2004

Evaluation of new methods for processing drilling data to determine the cause of changes in bit performance

Jaime Solano
Louisiana State University and Agricultural and Mechanical College, jsolan1@lsu.edu

EVALUATION OF NEW METHODS FOR PROCESSING DRILLING DATA TO DETERMINE THE CAUSE OF CHANGES IN BIT PERFORMANCE

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
Requirements for the degree of
Master of Science in Petroleum Engineering

in

The Department of Petroleum Engineering

by
Jaime Solano
B.S. C.E, Universidad Industrial de Santander, 1989
B.S. P.E, Universidad Industrial de Santander, 1992
May 2004

ACKNOWLEDGEMENTS

" Gracias Dios mío por hacer de mi la persona que tu quieres que yo sea "

I would like to thanks the LSU Craft and Hawkins Department of Petroleum Engineering. Especially, I express my appreciation to my major professor, Dr. John Rogers Smith, under whose supervision and guidance this research was conducted. Special thanks to my examining committee members, Dr. Christopher White, Dr.Yilmaz Karasulu, and Dr.Andrew Wojtanowicz, for their suggestions and cooperation.

I wish to thank my country, Colombia, because through the Universidad Industrial de Santander and ECOPETROL I received the academic knowledge and professional experience to be successful in my life. Special thanks to my manager in ECOPETROL, Dr. Alexis Meza. I could not have completed my master degree without his immeasurable help.

I am really thankful to the State of Louisiana (U.S.A), where my family and I received invaluable education.

I wish to thank my friends who assisted my family and me during my graduate studies. Special thanks to Orlando Uribe, Miguel Armenta, Christian Ferreira, Ben Sizemore, Nayibe Maxson, William Correa, Juan Velasco, Jorge Calvete, Jaime O. Castro, Orlando Pinto, Carlos Vega, Henry Arias, Liliana Sanchez, and Gloria Alfonso. Also, thanks to Jaime Becerra and Marcela Medina because you really cheer me up with your words when I need it. I cannot forget the unconditional assistance that my family received from my mother in law, Esperanza de Florez.

This work is specially dedicated to my family. My wife, Luz Edith, for her love, support and patience. My daughters, Diana, LuzA, and Laura, the reason of my life. My parents, Jaime and Mimi, my brother, Juan Carlos, and my sisters, Mary and Vicky, for their love and for encouraging me to complete my goals. Thanks to my grandmas, Olinda and Victoria, because wherever I am I feel their blessings.

TABLE OF CONTENTS

APPENDIX

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Drilling operations represent the major cost in finding and developing new petroleum reserves. Poor drilling performance when drilling deep shale and strong rocks, as evidenced by slow rate of penetration (ROP), has a significant detrimental impact on drilling costs. Also, it has been concluded that bit balling is the main cause of low ROP when drilling deep, clay-rich shale with water-based mud. In addition, it is estimated that a potential saving in drilling cost of hundreds of millions of dollars a year can be obtained if bit balling is mitigated and ROP is improved.

Several methods have been developed in order to improve bit performance. Recently, Arash Aghassi [1] and John R. Smith [37] proposed one of them, which uses simple drilling data to identify bit balling and lithology change as two separate effects through the calculation of five diagnostic parameters and comparing these values to a "baseline" zone. The objective of this research is to apply, evaluate, and improve the method proposed by Aghassi and Smith.

A set of down-hole well data and several sets of surface well data were used to evaluate the method. The diagnostic parameters of Aghassi's method were calculated, first using the down-hole data, and then with surface data. All of the results were correlated with and compared to wire-line logs. As a result, the utility of using surface data was confirmed. The overall utility of the method and its diagnostic parameters for detecting the occurrence of, and increases in, the severity of bit balling as distinguished from drilling into a stronger rock were evaluated. The results were very sensitive to the selection of the baseline; also, when drilling strong rock, the interpretation of the diagnostic parameters is sometimes that the bit is balled.

Statistical "Logistic Regression" models were developed and evaluated as a means to solve these problems. Those models were applied using several sets of well data. As a result, it was

determined that the logistic regression can potentially provide a basis for distinguish between bit balling and strong rock. It can be used independently, but it is more effective as a complement to Aghassi & Smith's method.

1. INTRODUCTION

1.1. General Description and Objective

The success in drilling operations depends on bit performance. Bit performance is usually defined in terms of the cost per foot drilled. It depends on many factors such as type of bit used, formation drilled, drilling fluid properties, depth of the well, bit tooth wear, bit hydraulics, and drilling operating conditions[15]. Bit performance is directly affected by the rate of penetration (ROP). Rate of penetration is the rate of forward progress of the bit measured in units of distance drilled per unit time. With a bit performing at high penetration rates, the drilling cost per foot is lowered. The penetration rate is often affected negatively by a phenomenon called "bit balling" which usually occurs when drilling clay-rich shale with water-based mud. Also, other factors such as a change in lithology can reduce the rate of penetration negatively affecting bit performance.

While a formation is drilled, bit balling occurs when rock-cuttings accumulate at and/or adhere to the bit, which causes a lower ROP than expected. Another situation that results in a lower ROP is when a formation change from a weak easily drillable formation to a hard rock is encountered. In the field, these two phenomena show similar behavior. Therefore, it is often difficult to distinguish which is happening. When a hard rock is encountered, an effective way to increase the ROP is to increase the WOB. However, if the situation is bit balling and the WOB is increased, the balling will be severe and irreversible[3]. The severity of bit balling could potentially be minimized by reducing WOB quickly after it has been detected. As a consequence, severe bit balling can occur in the field if the driller concludes that a reduced ROP is caused by encountering stronger rock when the cause is balling, but it could potentially be avoided if properly diagnosed.

A new methodology for diagnosing the cause of poor bit performance using drilling parameters measured at surface was introduced by Arash Aghassi[1]. This method determines the cause of changes in penetration rate by distinguishing bit balling from other causes, such as a change in lithology, through the calculation of five diagnostic parameters. Knowing the cause of a change in bit performance allows more appropriate actions by the driller to maximize bit performance under new drilling conditions.

The objective of this research is to apply, evaluate, and improve the method proposed by A. Aghassi. The method will be evaluated using both down-hole (MWD, measured while drilling) data and surface data. One of the goals is to determine whether the diagnostic parameters calculated using surface data respond in the same way as when calculated using more accurate and responsive down-hole data. Several data sets from different regions in the United States and Latin America will be used to determine the responsiveness and consistency of the diagnostic parameters for distinguishing balling from strong rock in wells with well-defined lithology.

1.2. Research Strategy / Plan

This study and the method developed by Aghassi[1] are part of a larger research project entitled "Automated Rig Controls for Improved Drilling Costs." The overall objective of this project is to provide a logical basis to diagnose bit performance using real-time drilling data in order to use the rig control system to make effective changes in operating parameters to achieve maximum bit performance. In this context, diagnosing bit performance means identifying the cause of a change in penetration rate, especially distinguishing bit balling from other causes of poor performance such as encountering a stronger rock.

The principal goal of this study is to develop a more "automated" approach for implementing Aghassi's method that could potentially be applied to real-time data. The plan for accomplishing that goal is:

1) To gain understanding of the topic by reviewing previous research.

2) To study, apply, and evaluate Aghassi's method using the down-hole and surface data from Matagorda Island # 6 well.

3) To evaluate the accuracy of the method to diagnose strong rock using surface drilling data and wire-line logs from 5 additional wells with well-defined lithology. The results of applying the diagnostic parameters of Aghassi's method to actual drilling data will be evaluated by comparing the diagnostic conclusions to detailed drilling records and logs of the wells studied. For example, indications of drilling strong rock will be compared to wire-line log data such as gamma ray, resistivity, porosity, and acoustic logs and to lithology descriptions of cuttings to determine the validity of the diagnosis. Likewise, indications of balling were compared to lithology based on log interpretation, as well as the subsequent bit performance history and visual inspection of the bit after the trip out of the hole.

4) To develop and apply logistic regression models in several wells, or intervals of a well, to evaluate these as alternative methods to distinguish bit balling and strong rock. Different inputs of data such as drilling parameters, conventional parameters, normalized parameters, and new diagnostic parameters will be tested.

5) To determine which parameters and model provide the most conclusive diagnosis for strong rock and balled bit independently.

8) Using logistic models as a complement to Aghassi's method in order to improve the reliability of the diagnosis. Determine how the logistic statistical model impacts the final diagnosis.

Recorded field data was acquired from Nabors Drilling, MD Totco, BP, H&P Drilling, BR, and two additional companies that have requested to be anonymous for these evaluations.

The first two chapters of this work are intended to introduce and explain the poor bit performance problem and the methods and models proposed to avoid and minimize its impact on field drilling operations.

Chapter 1 introduces the problem and provides an overview of the work performed in this research to address the objectives.

Chapter 2 provides a review of the existing technical literature about experimental methods and models, field experiences, causes of the problem, and proposed solutions for poor bit performance. In this chapter, the method proposed by Aghassi to detect the cause of a change in bit performance and distinguish bit balling from strong rock as two different causes of the problem is introduced and explained.

Chapter 3 describes the overall methodology used to evaluate the method developed by Aghassi. One example using down-hole data and one example with intervals of strong rock are shown to explain the results of the evaluation. The effect of selecting the location of the baseline zone was also studied and explained in this chapter. In addition, the possible causes of error are discussed. Finally, the reliability of the method is evaluated.

Chapter 4 introduces the concept of logistic regression, a statistical model proposed to diagnose bit performance. A parallel between linear regression models and logistic models is made. In addition, the reasons for and specific uses of logistic regression, the calculation of coefficients, as well as simple ways to evaluate the statistical significance of the coefficients and the effectiveness of the models are explained.

Chapter 5 includes examples about how to use logistic regression for diagnosing the situations of a balled bit and strong rock independently. Several models are developed using different independent parameters. The models and parameters are evaluated and the best models are used for final logistic diagnosis of strong rock and bit balling. Then, the methodology proposed for distinguishing between strong rock and bit balling by applying logistic regression is described.

Chapter 6 introduced the concept of using the logistic regression models as a complement to Aghassi and Smith's method in order to improve the efficiency of the diagnosis. The methodology proposed is illustrated with an example, utilizing a complete drilling data set of a well from Oklahoma. First, the Aghassi's method is applied and the reliability of its diagnosis is evaluated. Then, logistic regression models are developed, applied, and evaluated to diagnose specific situations of strong rock and bit balling. Finally, the effect of using a combination of Aghassi's method and the logistic models on the final diagnosis is evaluated.

Chapter 7 is a summary of the overall study with conclusions and recommendations for further research.

2. LITERATURE REVIEW

2.1. Overview

There have been many attempts to improve bit performance. Several author [3, 17, 18, 19, 20] developed models and methods to prevent or reduce bit balling. Some researchers[12, 13, 16] proposed methods to predict and optimize the rate of penetration. Others[21, 22, 23, 24] studied the effects of mud properties on rate of penetration. Falconer[4], Smith[2, 6, 7], and Pessier & Fear [8], developed and proposed several normalized and dimensionless parameters such as Specific Energy, Force Ratio, and Apparent Formation Strength, for characterizing and diagnosing bit performance and lithology changes.

2.2. Bit Balling

There are many possible causes of poor bit performance. When drilling deep, over pressured, shale formations, the causes can be bit balling, high confining pressure, and the presence of laminations of strong siltstone in the shale[6, 9]. However, bit balling is widely considered as the main cause of poor bit performance in shale[3,9,27], especially deep shale are being drilled with water-based mud [7,9].

Bit balling occurs when rock-cuttings accumulate under and/or adhere to the bit (face, cutters, and body), which causes a lower rate of penetration than expected. Historically, there has been confusion about what causes drilling cuttings to agglomerate and attach themselves to the bit and bottom-hole-assembly.

According to Ledgerwood[27], bit balling occurs because of mechanical and chemical factors. The mechanical explanation is that when shale material fails due to the cutting action of the bit, a sudden increase in formation porosity and a correspondent drop in pore pressure occur. This phenomenon is known as rock dilation. The chemical explanation is that the clay-rich shale exhibits a pronounced tendency to hydrate and the drilling fluids wet the surface of the bit. Consequently, due to the combination of these two effects, low pore pressure and the tendency of shale to hydrate, the cuttings tend to "vacuum" themselves on to the face and body of the bit[28].

Apart from reduction of penetration rate, bit balling can cause additional problems such as stuck pipe, high torque, lost circulation, increasing of time required for trip due to swabbing and tight hole, and difficulties when running the casing[1].

According to Bourgoyne[15], bit performance depends on many factors such as bit design, mud properties, bit operating parameters (weight-on-bit and rotary speed), and hydraulics. Several researchers have proposed solutions for low bit performance problems based on the study of these factors.

Many authors have studied the effect of mud composition on rate of penetration[3, 21, 22, 23, 24]. For example, Cheatham[3] made several tests in a full scale-drilling simulator with both water-based mud and oil-based mud. He concluded that balling occurred depending on: (1) mud type, (2) weight-on-bit, and (3) over pressure. He also concluded that balling ranged from slight to severe for the uninhibited water-based mud and did not occur for oil-based mud.

A study made by Smith[9] shows that in one area of the Gulf of Mexico, use of oil-based mud and bladed PDC bits resulted in three times faster rate of penetration and three times lower the cost per foot than similar bits in water-based mud. He estimated a potential world-wide saving of $500 million per year can be obtained if the bit performance in water-based mud is improved to be equivalent to bit performance in oil-based mud.

All researchers concluded that using oil-based mud (OBM) or synthetic-based mud (SBM) can inhibit the cohesion between cuttings and consequently reduce or avoid bit balling when drilling shale, especially in deep, high pressured wells. However, because of environmental regulations and economics, in many areas around the world the use of OBM and

SBM is unattractive. These areas are growing, and that has persuaded operators and service companies to invest considerable resources into trying to improve bit performance with water-based mud[29].

Proper selection and adjustment of bit operating parameters are critical to effective bit performance. Weight-on-bit (WOB) is the most important parameter to improve bit performance and avoid or minimize bit balling. According to Cheatham and Nahm[3], reducing WOB quickly enough after an incipient balling situation has been detected can prevent severe bit balling. Using a low, controlled WOB is advantageous in minimizing the effect of bit balling. Some reports indicate that knowing the down-hole WOB, from a measurement while drilling (MWD) tool, allows careful WOB control and improves of the average rate of penetration[6, 9].

Bit design is another opportunity for minimizing bit balling. Several authors and service companies have addressed this topic[9,15,30,31]. Tooth length; number of cutters; cutter exposure or blade standoff; size, shape, surface, and angle of the cutter; and nozzle and jet design are some of the many bit characteristics which affect ROP and bit performance. Recently, M. Rujhan Mat[29] proposed the application of a new low-friction flouropolymer-based coating applied to PDC bits to minimize and often prevent balling while drilling shale formations. This technique was first tested in laboratory and has also been tested successfully with several bit runs in the field, showing that the rate of penetration can be increased more than double in certain cases.

2.3. Estimation and Improvement of Penetration Rate

Several models and methods have been published for predicting, and therefore potentially optimizing penetration rate (ROP)[13,14,15,16,25], but apparently none is in regular use as a drilling planning or performance analysis tool. Current drilling engineering models cannot predict ROP for different combinations of bit design, rock and drilling fluid. That makes the optimization of ROP difficult to achieve or demonstrate.

Recently, M. Fear[12] developed a method that identifies which factors are controlling ROP in a particular field and group of bit runs. The method uses mud logging data, geological information, and drill bit characteristics, to produce numerical correlations between ROP and applied drilling parameters. These correlations are then used to generate recommendations for maximizing ROP in the subsequent drilling operations. The advantage of this approach is that, rather than use a model developed in the laboratory with relationships between ROP and drilling parameters, real data is studied to determine which factors are controlling the ROP. Perhaps the greatest contribution this method could make to analytical ROP modeling is to indicate with which variables a model would need to be constructed for each particular application.

2.4. Methods to Diagnose Bit Performance

Early detection of the on-set of bit balling occurs is very important to avoid irreversible and potentially severe bit-balling situation. Cheatham[3] determined that bit balling can become severe in a matter of fifteen seconds. He said when detected in this period of time, balling is largely reversible by decreasing weight-on-bit and washing the bit just off-bottom; otherwise, balling becomes irreversible. For these reasons, researchers have developed several models, methods, and parameters to diagnose bit balling and others situations that cause low bit performance.

2.4.1 I.G. Falconer [4], 1988

This method introduced two normalized parameters, dimensionless torque (T_D) and apparent formation strength (FORS) to diagnose and separate bit effects from lithology effects during drilling operations. "Dimensionless torque, T_D, is proportional to the bit efficiency and the ratio of the in-situ shear strength to the in-situ penetration strength." "Apparent formation strength, FORS, is proportional to the in-situ penetration strength of the rock and inversely proportional to the bit efficiency." T_D and FORS are defined by the following equations:

$$T_D = \frac{12 * DTorque}{DWOB * (Bit\ Diameter)} \quad \text{...(2.1)}$$

$$FORS = \frac{60 \times DWOB * RPM}{12 \times ROP * (Bit\ Diameter)} \quad \text{...(2.2)}$$

Where *DTorque* and *DWOB* are the down-hole torque and down-hole weight-on-bit respectively measured with MWD tools. According to the authors, the techniques can diagnose and provide information about lithological correlation (classified into three categories: porous, argillaceous (shaly), and tight, corresponding to high, medium and low torque respectively), wear state of the bit teeth only in shale, and excessive torque and cone locking.

2.4.2 C.A Cheatham[3], 1990

Using a full-scale drilling simulator the author determined that the ratio of bit torque to weight-on-bit is a reliable indicator of the degree of bit balling. At the beginning of balling, the ratio increased above that of clean bit and was erratic. For the situation of severe bit balling, the ratio was lower than that of a clean bit and more erratic.

2.4.3. J. R. Smith[6, 7, 9, 10], 1998-2000

J.R. Smith identified the symptoms of low bit performance during several bit runs in the field and matched them to the symptoms resulting from different possible causes in controlled laboratory tests. He primary used two measures for evaluating bit performance: Specific energy (E_S) and force ratio (R_f). Specific energy is mechanical work being done at the bit per unit volume of rock removed. The force ratio is the ratio of the force acting to push the bit tooth or cutter laterally through the rock to break and remove it divided by the force acting downward to engage the tooth or cutter in the rock. This normalized parameter is similar to the dimensionless torque defined by Falconer [4]. E_s and R_f are defined by the following equations:

$$E_S = \frac{WOB}{(Borehole\ Area)} + \frac{120\pi * RPM * Torque}{(Borehole\ Area) * ROP} \quad \text{..................................(2.3)}$$

$$R_f = 48\frac{Torque}{(Bit\ Diameter) * WOB} \quad \text{..(2.4)}$$

The equation for E_s was taken from Pessier and Fear[8]. The coefficient shown as 48 in the equation for R_f applies to bladed PDC bits and is replaced with 36 for roller cone bits, body-set PDC bits, and conventional diamond bits.

Smith concluded that a lower force ratio and a higher specific energy than expected in shale are quantitative symptoms of bit balling situations. In addition, he observed that strong siltstones also cause quantitative symptoms similar to the slow drilling shale problems. However, nothing caused specific energy as high as recorded when bit balling occurred. These research studies were the basis for Aghassi's method for distinguishing bit balling and strong rock as two separate causes of poor bit performance.

2.5. Aghassi and Smith's Method[1,37], 2002

In order to improve bit performance, Arash Aghassi and J. R. Smith [1,37] proposed to use simple drilling data to develop a method to identify bit balling and lithology change as two separate effects through the calculation of the following diagnostic parameters:

- ROP/WOB: The value of this diagnostic parameter is similar to 1/FORS as defined by Falconer [4, 5]
- Torque/ROP: The value of this diagnostic parameter is similar to Specific Energy concepts from other research studies [6, 7, 8]
- Torque/WOB: The value of this diagnostic parameter is similar to other research studies parameters, such as Force Ratio [3,6 7, 9, 10]

- F (TORQUE, WOB): New diagnostic derivative parameter as function of Torque and WOB, which was introduced by Aghassi and Smith [1,37]
- G (ROP, WOB): New diagnostic derivative parameter as function of rate of penetration and WOB, which was introduced by Aghassi and Smith [1,37]

The definitions of "F" and "G" are not shown in this study because the originators, A. Aghassi and J. R. Smith, are protecting the intellectual property for these new derivative parameters.

Applying these five diagnostic parameters, the method requires only three drilling parameters, WOB, TORQUE, and ROP. This method and its diagnostic parameters were primarily developed using laboratory tests and then tested with real drilling data measured at the surface in one well.

Because the first three parameters (ROP/WOB, Torque/ROP, and Torque/WOB) are equivalent to those developed by previous researchers, they are described, in this study, as "Conventional diagnostic parameters". On the other hand, the parameters F (ROP,WOB) and G(Torque, WOB) introduced by Aghassi and Smith, are designated as "New Diagnostic Parameters".

In this method, all the values of the diagnostic parameters are compared to "baseline" values, which are located in an interval of relatively high ROP over a long shale section. Because "relatively high" is a subjective definition and a high ROP may not be achieved in a particular bit run, the selection of the baseline zone is essentially arbitrary. Once the baseline is defined and the diagnostic parameters calculated, compared and interpreted, it is possible to detect changes in bit performance and to determine whether the cause is a formation change or bit balling. The interpretation of the diagnostic parameters are summarized in Table 2.1 taken from Aghassi's thesis [1]

Table 2.1 – Summary of Observed Field Response of Diagnostic Parameters Compared to a Shale Baseline per Aghassi[1]					
First Diagnostic parameter Group				**Second Diagnostic parameter Group**	
Situations	**Torque/WOB**	**Torque/ROP**	**ROP/WOB**	**F**	**G**
Shale (baseline)	In Baseline	In Baseline	In Baseline	In Baseline	In Baseline
Sand	Larger	Varies	Larger	Larger Negative	Larger Negative
Shale (Cleaner Bit)	Larger	Same	Somewhat Larger	Larger Negative	Larger Negative
Shale (Severe Balling)	Slightly Smaller	Larger	Slightly Smaller	Much Smaller Negative	Noisy / Inconclusive
Stronger Rock	Slightly Larger or Close	Larger	Slightly Smaller or Close	Same or Slightly Larger Negative	Noisy / Inconclusive

In order to get more accurate results, the method corrects the rate of penetration using Lubinski's equation[11]. The reason for this correction is that conventional rate of penetration instrumentation does not provide a correct measurement to indicate the progress of the bit at the bottom of the hole[1, 11.] . In other words, the rate of penetration recorded in the field is almost always measured by movement of the top of the drill string at the surface rather than actual movement of the bit at the bottom of the hole. Consequently, Lubinsky[11] proposed an improved approach for estimating actual rate of penetration, ROP, at the bit by taking into account the theoretical drill string elasticity and its change of length caused by the effect of changing weight-on-bit, WOB, by using the following equations:

$$ROP = \frac{dD_s}{dTime} - K\frac{dWOB}{dTime} \quad \text{...(2.5)}$$

$$K = \frac{L}{144EA} \quad \text{..(2.6)}$$

More details about how to apply Lubinsky equations can be found in the appendix II of Aghassi's thesis[1].

The potential procedure designed by Aghassi[1] to detect the causes of changes and improve bit performance, especially detecting the onset of, or an increase in the severity of, balling is:

1) Find the baseline value for all diagnostic parameters using data from a previously drilled interval with the best ROP in shale within the bit run. Shale intervals may be inferred directly from small negative values of the derivative parameter "F". Shale intervals may also be implied by correlation to offsets or confirmed by cuttings.

2) Make sure other important drilling parameters, e.g., RPM, flow rate, pump pressure, mud properties, etc., are relatively constant.

3) If any of the diagnostic parameters change, use Table 2.1 to determine the probable cause so that the appropriate action may be recommended for the situation. Table 2.1 is based on trends in the diagnostic parameters observed in field and laboratory data.

4.) Take the appropriate action. For example, reduce the WOB if balling is becoming more severe, or increase the WOB if the ROP is low in a strong rock.

5) Update the baseline as needed for change in depth or drilling conditions such as RPM, flow rate, the bit, etc.

6) Repeat steps 2) through 5) as required throughout the bit run.

2.6 The Logistic Regression

Statistical binary logistic regression models are proposed in this research as an alternative way to diagnose bit performance, taking into account as "independent variables" the drilling and normalized parameters such as ROP, WOB, torque, specific energy (ES), apparent formation strength (FORS), force ratio (R_f) and the new parameters "F" and "G" proposed by Aghassi and Smith [1,37].

The Logistic Regression Model[33,34,36] is defined as:

$$\mathrm{Ln}\,[P / (1- P)] = a + B_1 X_1 + B_2X_2 + \ldots B_nX_n + e \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2.7)$$

or

$$[P/ (1- P)] = \exp a \ \exp B_1 X_1 \ \exp B_2X_{2\ldots} \ \exp B_nX_n \quad \exp e. \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2.8)$$

Ln is the natural logarithm, $\log_{exp}$, and exp=2.71828… P is the probability that an event occurs. a is the constant term, B_1 to B_n are the coefficient(s) on the independent variable(s), X_1 to X_n are the independent variable(s), and **e** is the error term.

The estimated probability is:

$$P = 1/[1 + \exp(-a - B_1 X_1 - B_2X_2 - \ldots B_nX_n)] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2.9)$$

In this regression, the outcome (probability that an event occurs) is measured on a binary scale between zero and one. Then, if the result is one or close to one, the outcome (e.g. bit balling) is true. On the other hand, if the result is zero or close to zero, the outcome is false. This technique allows the use of any number of independent variables or parameters for developing the model, but for every independent variable it is necessary to calculate its respective coefficient. In order to calculate the coefficients, this method requires a training data set from an interval with known occurrences of the events of interest, such as bit balling, strong rock, and routine drilling with correlative drilling data.

Given all the attention that writers and researchers have given to logistic modeling over the last 20 years, it is possible to find several instructive books and papers about the application of this statistical model in different sciences. F.C. Pampel[33], introduces and explain the logistic regression and the interpretation of its coefficients with elementary language and simple examples. Also, other sources such as Jansen[35] and Dobson[36] were studied and applied to develop this research. Professor Scott Menard [34] fully explains the estimation, interpretation, and diagnosis of logistic regression models. He also discusses the current computer software available for logistic regression. Logistic models can be found in several statistical computer packages. In this research, the "R" computer software was applied for calculation of the coefficients of the logistic models. This software is available from GNU General Public License, 1991, Free Software Foundation, Inc.

3. EVALUATION OF AGHASSI AND SMITH'S METHOD

The method developed by Aghassi and Smith[1,37] described in Chapter 2 (section 2.5) uses conventional and specific new diagnostic parameters to identify bit balling and lithology change as two separate effects. This method and its diagnostic parameters were primarily developed using laboratory tests and then tested with real drilling data measured at surface in a well. Three simple drilling parameters (Torque, WOB, and ROP) are used to calculate the conventional and new diagnostic parameters. These drilling parameters have different values when they are recorded or measured at surface as compared to measurements taken at the bit. Aghassi corrected the effect on ROP using the Lubinski's method for calculating ROP at the bit. Bit torque is typically significantly less than surface torque due to friction in the hole opposing rotation of the drill string. Aghassi [1] studied several simple approaches to treating the frictional torque as an offset and subtracting it from total surface torque to estimate bit torque, but he did not get satisfactory results. In addition, frictional drag also affects down-hole WOB, especially in directional wells.

One of the objectives of my study is to apply and evaluate the method using both down-hole MWD data and surface data from a well with known lithology. Comparing these results, it is possible to evaluate the accuracy and utility of the method when using surface data instead of data measured at the bit.

3.1. Application of the Method Using Down-hole Data

In order to determine if the diagnostic parameters calculated using "surface" data respond in the same way as when calculated using more accurate and responsive "down-hole" data, Aghassi & Smith's method was applied and evaluated using a set of data from the Matagorda Island Well # 6 (Interval 12,200'-12, 900', bit run # 12), where torque and WOB were recorded at both the surface and down-hole.

3.1.1 Data Description

Figure 3.1 shows the data from the interval of 12,200' – 12,900' in this well. The data includes: Rate of penetration (ROP), MWD down-hole weight-on-bit (DWOB), MWD down-hole torque (DTORQUE), logging-while-drilling gamma ray (DGR), surface weight-on-bit (SWOB), surface torque (STORQUE), and GR, resistivity (ILD, ILS, RAD, RPS), SP, sonic (DTLN, DTLF), and side-wall core analysis from wire-line runs.

The interval was drilled using a 10 5/8" HC-AR554G bit with four 16/32" jets. After drilling 700 feet, the bit was evaluated using the dull grading system as 3-3-BT-A-X-I-CT-TD. The bottom-hole-assembly had 410 feet of 8" drill-collars and 371 feet of 5" Hevi-Wate drill pipe. The average operating parameters were 100-130 rpm, 680-700 gpm flow rate, 17.7 ppg water-based mud, and 3050 psi pump pressure.

From figure 3.1, it can be observed that although there was often a quantitative difference between "down-hole" and "surface" measurements of torque and weight-on-bit, they have almost identical trends versus depth. However, there are intervals where the quantitative values of surface WOB are significantly larger than down-hole WOB.

According to the interpretation of GR (LWD), SP, and sonic logs, it can be determined that the interval from 12,200' to 12,900' is mainly shale with a laminated sandstone and shale sequence from 12,270' to 12,300'. GR has a high, almost constant, value of 100 units in the shale section, and it decreased to 50 units in clean sands. The sonic log also clearly differentiated the shale and sand zones. The shale sections have average sonic values of 130-140 usec/ft and are apparently weak. The sand sections have lower values of travel time of about 75 usec/ft and are evidently stronger. This interpretation is ratified by the sidewall core analysis taken from the interval 12,288'- 12,300', which were described as sandstone with porosity around 20%. In addition, samples taken from the interval 12,267'-12,307' were described as shale.

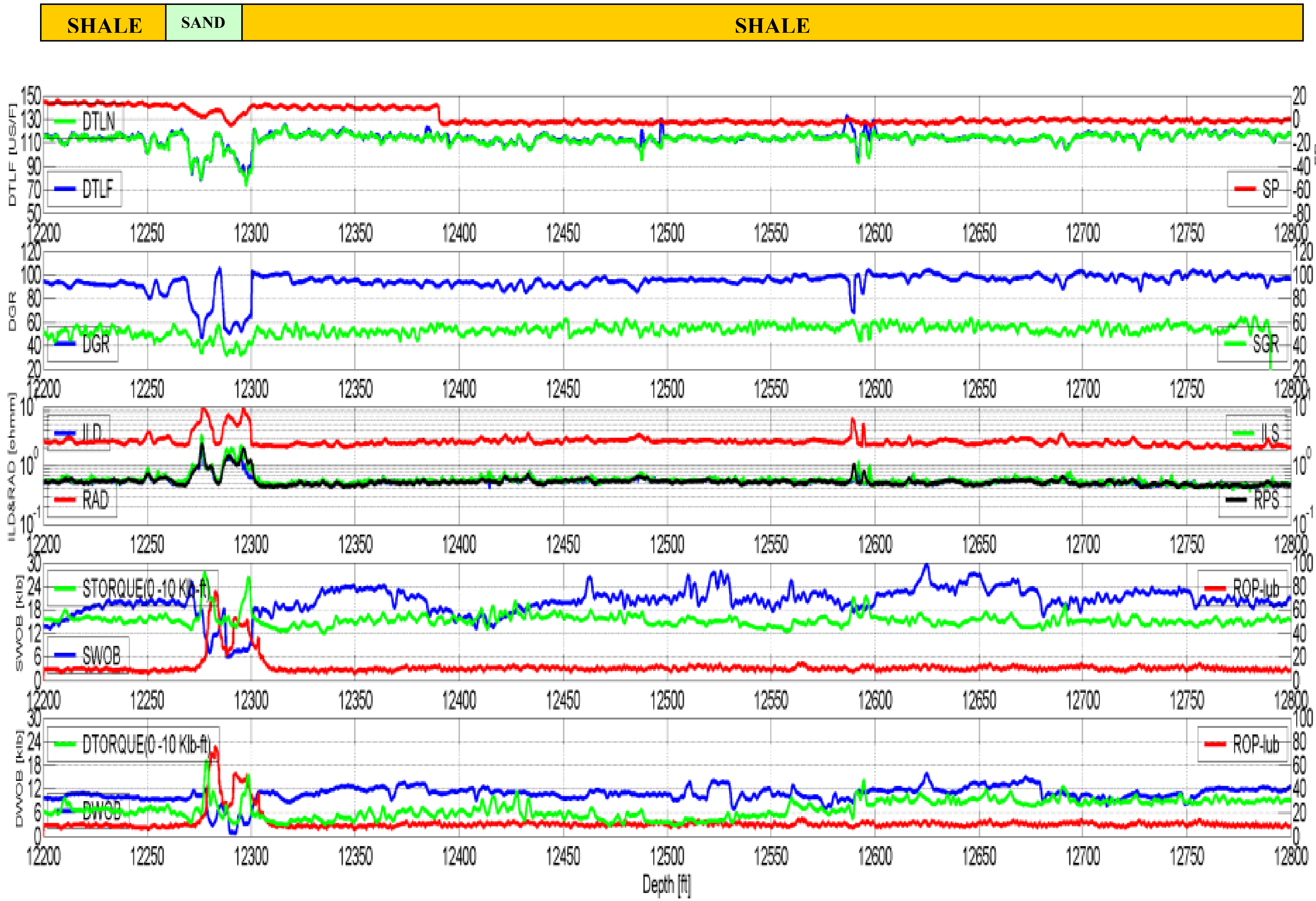

Figure 3.1. Matagorda Island Well # 6 Data

This interpretation is ratified by the sidewall core analysis made in laboratory, which shows sandstone with porosity around 20% between 12,288' and 12,300', and 100% shale at 12,307'.

Taking into account this lithological interpretation and the behavior of the drilling parameters shown in Figure 3.1, it can be determined that in the long shale intervals at 12,200'-12,270' and 12,300'- 12,900', the ROP is almost constant at about 10 ft/hr and the other parameters such as WOB and torque show only a slight variation. For this reason, the application and evaluation of Aghassi and Smith's method will be focused on the interval from 12,250' to 12,380 where lithology changes between sandstone and shale are associated with considerable variation of the drilling parameters.

The data from Matagorda Island Well # 6 was only available on logs plotted versus depth. There was no data available versus time. Although Aghassi's method is basically for working with high-resolution data versus time, the method was applied and validated with this data by converting the depth-based data to time-based data. Plots were digitized at 0.1 ft intervals and depths were converted to times dividing by the average value of ROP.

3.1.2 Application and Evaluation of the Method (Interval 12,250'-12,380')

Figure 3.2 shows the behavior of GR (LWD) log and the three drilling parameters: WOB, Torque and ROP measured at surface and down-hole (MWD) for the interval 12,250' – 12,380'. The ROP is corrected by Lubinsky's equation as recommended by Aghassi [1, 11].

Also, figure 3.2 shows in detail a lithology interpretation for this interval. From 12,250' to 12,270' there is a shale zone. Then, a shaly sand (SH-S) interval is drilled from 12,270' to 12,275', followed by a clean sand (S) zone (12,275'-12,278'). After that, the sand becomes shaly

(SH-S) in the interval 12,278'-12, 283', and then a short shale (SH) interval, called the intermediate shale herein, is encountered from 12,283' to 12'287. The GR log shows this intermediate shale clearly. After this intermediate shale, a longer sand interval from 12,287' to 12,300' is drilled. Finally, from 12,300' to 12,380' a shale zone was drilled.

Taking into account the lithology and the behavior of drilling parameters (ROP, WOB, Torque), the following operational interpretation can be made. The bit run was started at 12,200', see figure 3.1, with a relatively high value of WOB (15-18 klbs. at surface and 9-10 klbs. on down-hole), getting a ROP of 9 ft/hr which, according to Smith [9], is considered a low ROP because it is lower than 25 ft/hr. For this reason, there is the possibility that the bit was balled from the beginning of the run. Apparently, when the bit encountered the first sand from 12,270' to 12,283', it began to clean up, and the ROP began increasing significantly below 12,275', reaching a maximum value of 76 ft/hr with a relatively low WOB (3 klbs on down-hole) at about 12,283'. After that, the intermediate shale (12,283' – 12,287') was drilled with a down-hole WOB increasing to 7 klbs, and the ROP decreased to about 27 ft/hr. Then, a second sand at 12,287'-12,300' was drilled, and the ROP began to increase until a value of 50 ft/hr was achieved with a low DWOB of 2-5 klbs. Finally, a very long shale interval from 12,300' to 12,900' was drilled; there, the ROP decreased drastically, reaching a value of 8 ft/hr at 12,310'. Although they tried to improve the drilling performance increasing the SWOB from 6 klbs at 12,297' to 20 klbs at 12,310' and 24 klbs at 12'345, the ROP remained low at 8-10 ft/hr until the end of the run. Apparently this happened because the bit was completely balled again at 12,310'. Knowing the lithology and the large influence it had on ROP, this interval was analyzed further to compare results of using the proposed method using surface and down-hole data.

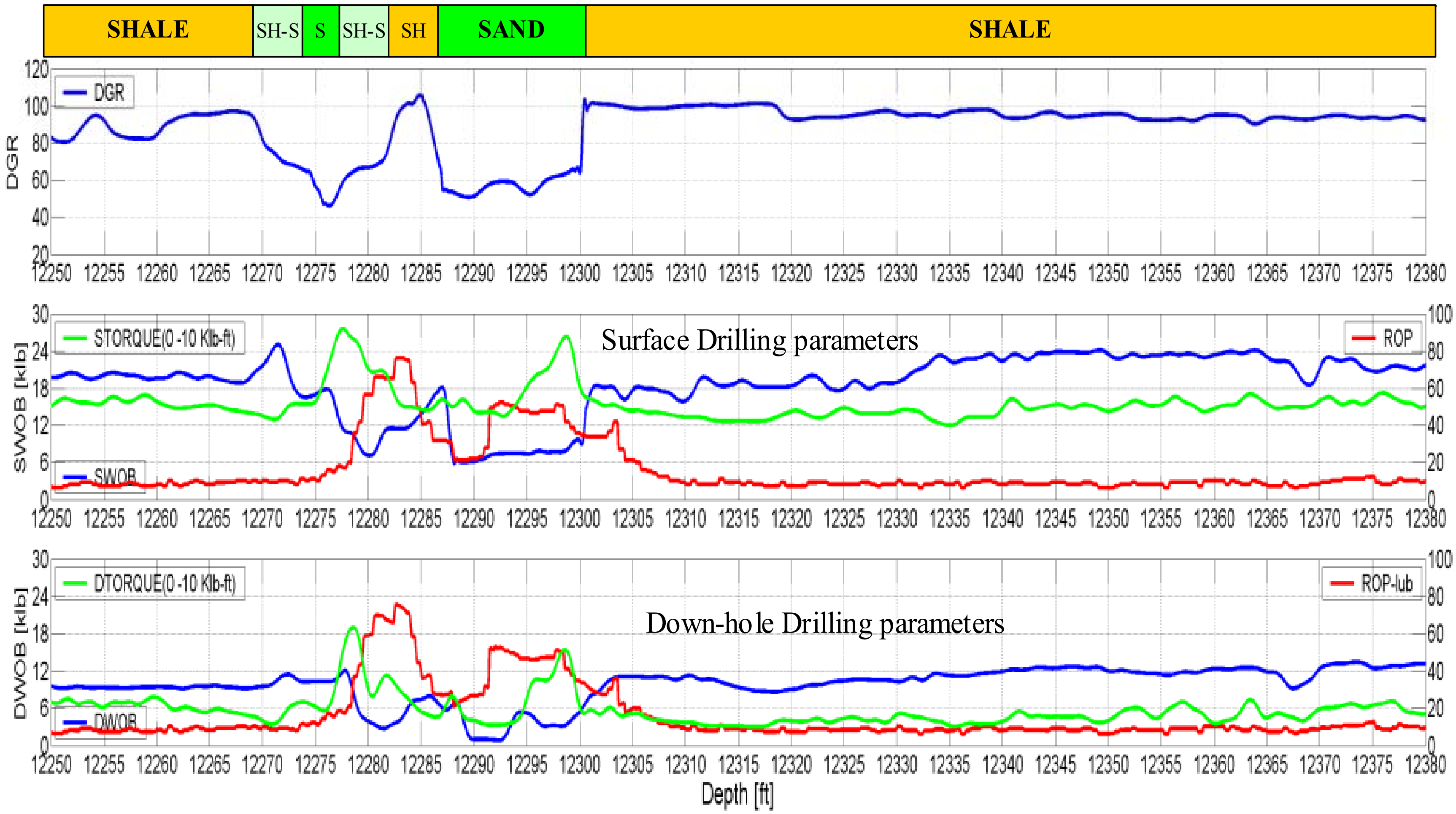

Figure 3.2 Geological Interpretation from Wire-line Logs and Drilling Parameters (12,250'-12,380')

3.1.3 Selection of the Baseline

The first critical step when applying Aghassi and Smith's method is to determine the "baseline" zone, which is preferably located in an interval of relatively high ROP over a long shale section. For this bit run, the long shale sections (12,200'-12,270' and 12,300'-12,900') do not have a relatively high ROP. Moreover, the average ROP (10 ft/hr) is considered low. On the other hand, the intermediate shale (12,283'-12,287') was drilled with a high ROP (70 ft/hr declining to 27 ft/hr), but this interval is short and operating conditions are not consistent. As a consequence, there is not an obvious zone for locating a baseline.

I consider that there are two zones where the baseline can be located in the interval 12,250'-12,380'. One of them would be in an interval, 12,250'-12,260' for example, of the initial long shale section where there is low ROP of 10 ft/hr. Another zone is the short intermediate shale section, 12,283'-12,287', where there is a high ROP of up to 70 ft/hr. Then, the analysis of this section was made considering both baseline intervals. As a consequence, the diagnostic parameters calculated for the zone 12,250'-12,283' are compared with the values at the first baseline (12,250'-12,260'). After that, the second baseline (12,283'-12,287') is defined, and all values of the diagnostic parameters from there to the end of the run are compared with the values at this second base line.

3.1.4 Calculation and Evaluation of the Diagnostic Parameters

The diagnostic parameters of Aghassi and Smith's method were calculated, first using the down-hole data, and then they were calculated with surface data and the values compared. The results were matched and validated versus the lithology interpreted from wire-line logs. The results are shown graphically in Figures 3.3 and 3.4.

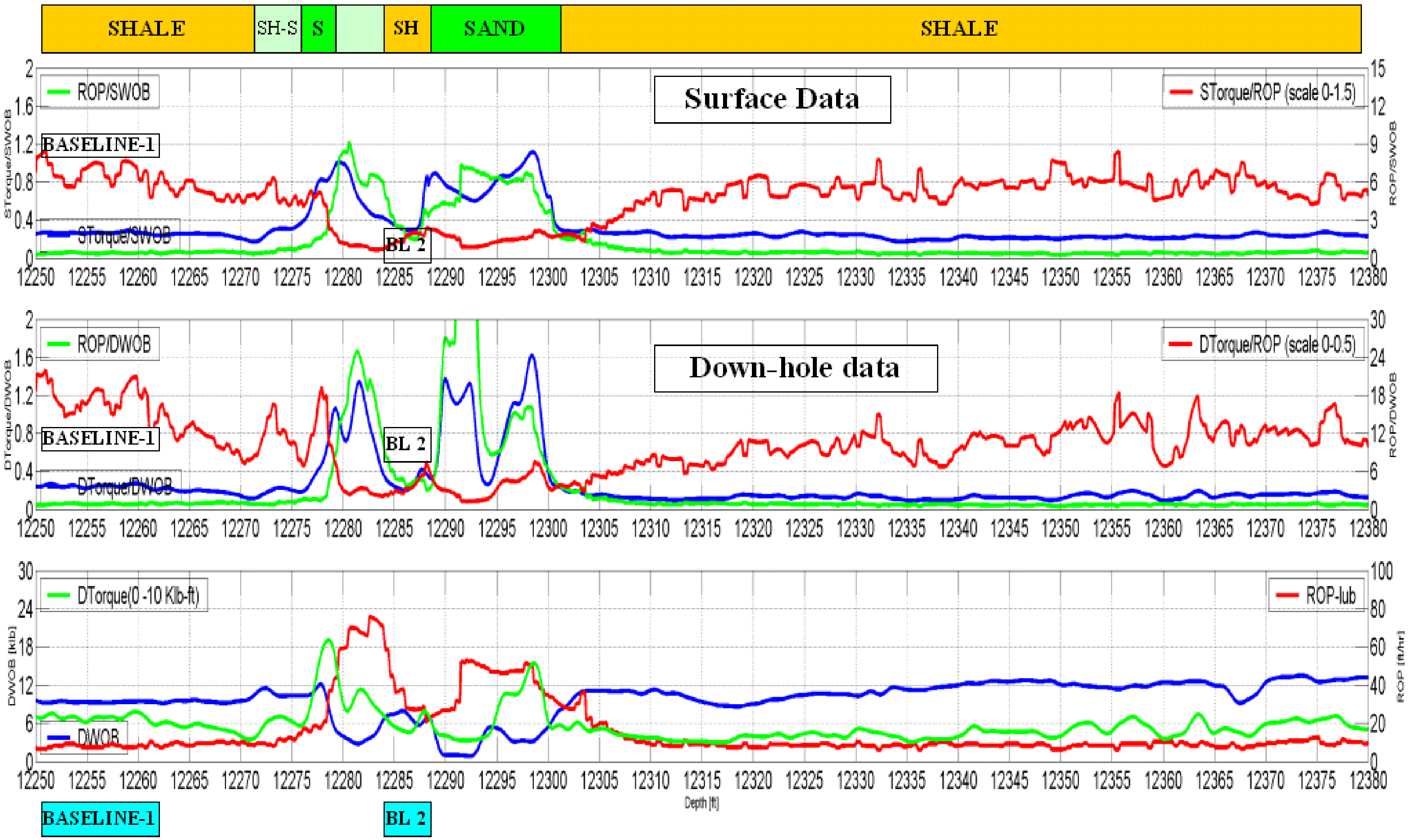

Figure 3.3 "Conventional" Diagnostic Parameters Using Down-hole and Surface Data for Matagorda Island Well # 6

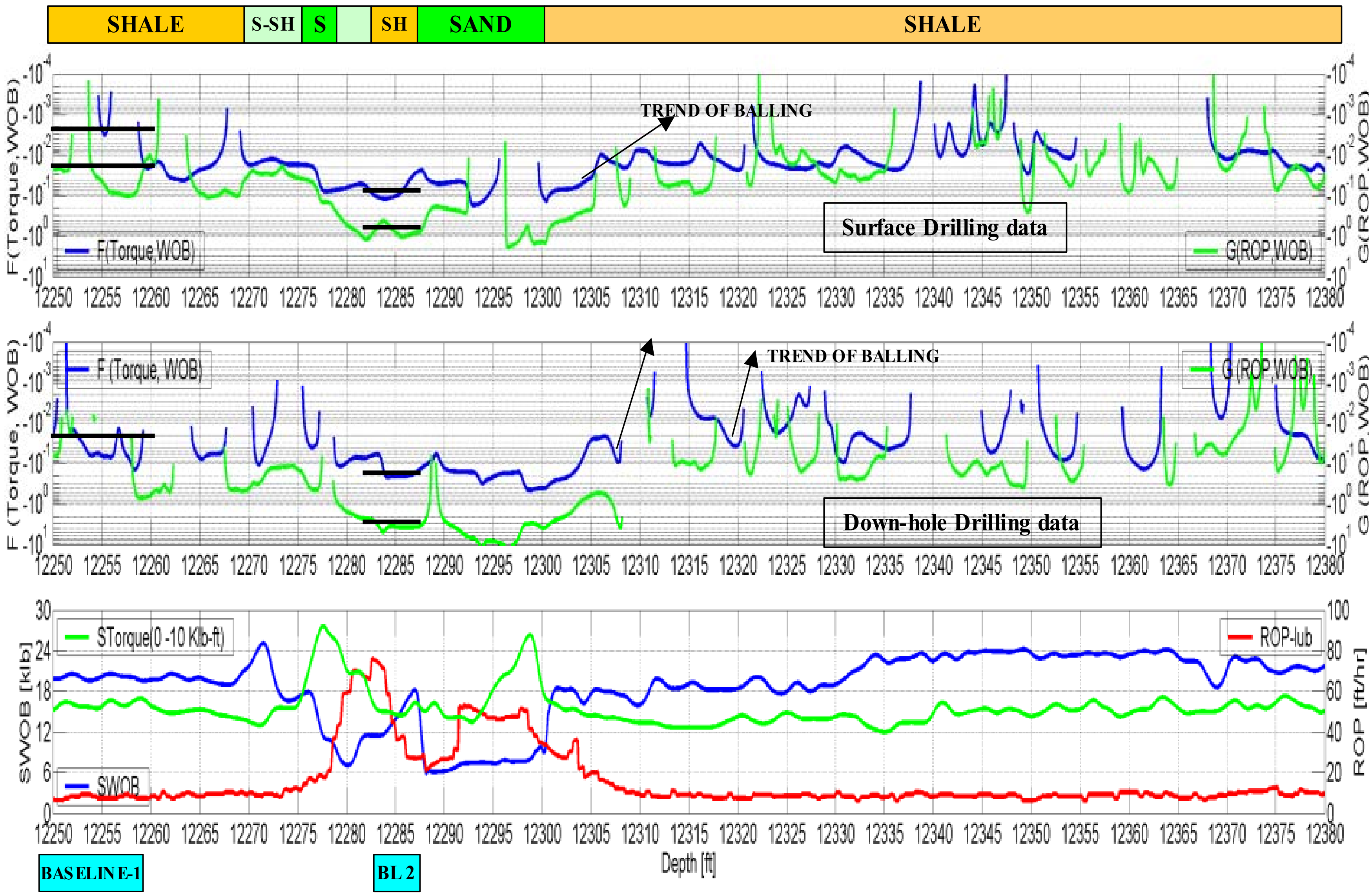

Figure 3.4 “New” Diagnostic Parameters using Down-hole and Surface Data for Matagorda Island Well # 6

Figure 3.3 shows the down-hole drilling parameters (DWOB, Dtorque, ROP-lub), the location of the baselines, and the plots of the "conventional diagnostic parameters" versus depth for surface and down-hole data. Figure 3.4 shows the surface drilling parameters (SWOB, Storque, ROP-lub), the location of the baselines and the logarithmic plots of the "new diagnostic parameters" versus depth calculated with surface and down-hole data. At some intervals of these plots, the F and G curves are truncated because these parameters took values very close to zero or even positives, which are out of the scale of the graphic. Tables 3.1 to 3.6 show the interpretation, for each interval analyzed, of the diagnostic parameters when using surface and down-hole respectively.

Once the first baseline is defined, the diagnostic parameters are calculated every 0.1 foot. A linear regression of 50 data points is used to define the derivative parameters F and G and the dWOB/dt needed for Lubinsky equation. When all the calculations are done, the diagnostic parameters are evaluated for every interval shown in Figure 3.2 as follows:

- 12,260'-12,270': In this interval of shale, the conventional diagnostic parameters Torque/WOB and ROP/WOB have the same value as the first baseline interval, baseline-1, for either surface data or down-hole data, see Figure 3.3 and Table 3.1. The conventional parameter, Torque/ROP, is decreasing compared with its value at the baseline-1 zone, which occurs because the value of Torque is decreasing slightly while the ROP keeps a constant low value of 10 ft/hr in this interval. The new diagnostic parameters, F and G, have erratic values in the baseline-1 zone, but it seems that for surface data the F parameter has a larger negative value than its value at the baseline-1; in contrast, for down-hole data, it has a smaller negative value. The G parameter seems to have the same value as its value in the baseline-1 zone, see Figure 3.4. In conclusion, three of the five diagnostic parameters have the same value as their values in the baseline-1, and the indications from the other two parameters are inconclusive, so the most likely situation is that the interval 12,260'-12,270' has the same lithology, shale, as the baseline-1 interval, which is the same interpretation made from wire-line logs, see DGR curve in Figure 3.2.

Table 3.1 Interpretation of diagnostic parameters for interval 12'260'-12,270'

Data measured at	Conventional Diagnostic parameters			New Diagnostic parameter		Interpretation	Is the Interpretation correct?
	Torque/WOB	Torque/ROP	ROP/WOB	F (Torque, WOB)	G (ROP, WOB)		
Down-hole	Same	Lower	Same	Smaller Negative	Same	Same as shale baseline-1	**Yes**
Surface	Same	Lower	Same	Larger Negative	Same	Same as shale baseline-1	**Yes**

- 12,270'-12,275': The lithology interpretation from log data says that this interval is shaly sand. There are variations in WOB and Torque, and the ROP shows a slight improvement toward the end of the interval. The behavior of the conventional diagnostic parameters is almost the same as for the previous interval; see Table 3.2, except that the value of Torque/ROP increases toward its baseline value but still lower than it. In addition, the new diagnostic parameter G has a smaller negative value, which has no conclusive meaning. The parameter F shows larger slightly negative values with surface data but it is noisy with down-hole data. In summary, there is no conclusive change in this interval, probably because the bit was balled from the beginning of the bit run and remains balled in this interval; as a consequence, the model does not provide a clear diagnosis.

Table 3.2 Interpretation of diagnostic parameters for interval 12'270'-12,275'

Data measured at	Conventional Diagnostic parameters			New Diagnostic parameter		Interpretation	Is the Interpretation correct?
	Torque/WOB	Torque/ROP	ROP/WOB	F (Torque, WOB)	G (ROP, WOB)		
Down-hole	Same	Lower	Same	Noisy	Smaller Negative	NO clear diagnosis	**NO**
Surface	Same	Lower	Same	Larger Negative	Smaller Negative	NO clear diagnosis	**NO**

- 12,275'-12,278': According to the geological interpretation, based on logs and core analysis, this interval is sand. In addition, the basic drilling parameters show a considerable improvement in bit performance. Rate of penetration increased from 10 ft/hr to 20 ft/hr, Torque increased and WOB was decreased. These are qualitative symptoms typical of a clean bit drilling in sand or another readily drillable formation. From Figures 3.3 and 3.4, it can be determined that the values of two conventional diagnostic parameters, Torque/WOB and then ROP/WOB, begin rapidly increasing versus the baseline-1 zone within this interval when calculated with either surface or down-hole data. This is a typical behavior of sand lithology or drilling with a cleaner bit; see Table 2.1. The new diagnostic parameters F and G calculated with surface data help to confirm the sand diagnosis of this interval, because they have larger negatives values than their values in the baseline-1, see Table 3.3. Consequently, the diagnosis from the diagnostic parameters of the Aghassi and Smith's method shows evidence that the bit is cleaning up or drilling sand in this interval, which agrees with the sand diagnosis from logging data and sidewall samples.

Table 3.3 Interpretation of diagnostic parameters for interval 12'275'-12,278'

Data measured at	**Conventional Diagnostic parameters**			**New Diagnostic parameter**		**Interpretation**	**Is the Interpretation correct?**
	Torque/WOB	**Torque/ROP**	**ROP/WOB**	**F (Torque, WOB)**	**G (ROP, WOB)**		
Down-hole	Larger	Lower	Larger	Noisy	Noisy	NOT clear from New par.	**Yes, only for conventional**
Surface	Larger	Lower	Larger	Larger Negative	Larger Negative	Sand or clean bit in shale	**YES**

- 12,278'-12,283': Although in this interval the rate of penetration continues increasing until the highest value of 70 ft/hr, the GR begins to increase indicating the section is becoming more shaly. Therefore, the geological interpretation of this interval is shaly-sand. In this interval, when comparing to the baseline-1 values, the conventional diagnostic parameters show the same results as for the previous interval (12,275'-12,278') for either surface or down-hole data. However, the parameters Torque/WOB and ROP/WOB are decreasing, see Figure 3.3 and Table 3.4, which is a trend back toward the baseline implying the presence of shale, and the final diagnosis tend to be "shale with clean bit". Moreover, F and G have larger negatives values, comparing to their values in baseline-1, for both surface and down-hole data. This is also an indication of drilling sand or shale with clean bit. In summary, the diagnosis from Aghassi and Smith's method is that this interval is something between sand and shale with clean bit, which is close to the geological interpretation made from logging data.

Table 3.4 Interpretation of diagnostic parameters for interval 12'278'-12,283'

Data measured at	Conventional Diagnostic parameters			New Diagnostic parameter		Interpretation	Is the Interpretation correct?
	Torque/WOB	Torque/ROP	ROP/WOB	F (Torque, WOB)	G (ROP, WOB)		
Down-hole	Larger	Lower	Somewhat Larger	Slightly Larger Negative	Larger Negative	Sand or clean bit in shale	**YES**
Surface	Larger	Lower	Somewhat Larger	Slightly Larger Negative	Larger Negative	Sand or clean bit in shale	**YES**

- 12,283'-12,287' (Baseline-2): Because this analysis was done long after the actual operation and with known lithology from logs, this interval was identified as a shale with relatively high ROP. For this reason, the baseline-2 is located in this short shale interval which is well defined by the GR log, see Figure 3.2. In this interval, the ROP varies from 70 Ft/hr to 27 ft/hr, which is considerably higher than the almost constant 10 ft/hr value in baseline-1. WOB varies from 12 klbs to 15 klbs at surface and from 4 klbs to 8 klbs at the bit; these ranges are lower than the 20 klbs at surface and 10 klbs at the bit in baseline-1. The average torque values are almost equal in the two baselines. In summary, bit performance in baseline-2 is better than in baseline-1, and that is probably because the bit was balled when drilling the shale section of baseline-1, then it was cleaned by drilling the sand sections before entered the shale section of baseline-2.

- 12,287'-12,300': GR response decreased from 100 to 50 units at the beginning of the interval, after that, the GR response is constant at about 50 until the end of the interval, see Figure 3.2. This shows clearly that this is a section of clean sand and is confirmed by the core samples taken from this interval. Comparing the values of conventional diagnostic parameters (Torque/WOB: much larger, Torque/ROP: little smaller, ROP/WOB: much larger) to their values at the baseline-2 zone, the diagnosis is also that this is probably an interval of sand, because according to Table 2.1, the increment of the diagnostic parameters Torque/WOB and ROP/WOB is a characteristic of drilling sand. Parameters F and G, calculated with down-hole data, have a somewhat larger negative value compared to their baseline-2 values, which is an indication that the formation drilled is sand or shale with clean drilling. The values of F and G calculated with surface data tend to be smaller negatives, see Figure 3.4 and Table 3.5, which is not conclusive in this case.

Table 3.5 Interpretation of diagnostic parameters for interval 12'287'-12,300'

Data measured at	Conventional Diagnostic parameters			New Diagnostic parameter		Interpretation	Is the Interpretation correct?
	Torque/WOB	Torque/ROP	ROP/WOB	F (Torque, WOB)	G (ROP, WOB)		
Down-hole	Larger	Lower	Larger	Slightly Larger Negative	Slightly Larger Negative	Sand	**YES**
Surface	Larger	Lower	Larger	Slightly Smaller Negative	Smaller Negative	NOT clear for new par.	**NO**

- 12,300'-12,350': The GR shows a drastic change at 12,300' increasing from 50 to 120, which is an indication of lithology change from sand to shale. After that, the GR response is constant at about 120 until the end of the bit run, see Figure 3.2. The drilling parameters also changed drastically. The penetration rate drops rapidly from 50 to 10 ft/hr over a distance of about 10 ft. Torque decreased, and WOB was increased at surface from 6 klbs to 24 klbs trying to improve bit performance, but the result was opposite. All these signals are qualitative symptoms of a bit-balling situation. The diagnosis from the conventional diagnostic parameters is that the bit is balled. Torque/WOB and ROP/WOB are lower while Torque/ROP is larger than their values in the baseline-2, see Figures 3.3, 3.4 and Table 3.6. This diagnosis is equivalent to high specific energy and low force ratio defined as symptoms of bit balling by Smith [6, 7, 9]. In addition, the new diagnostic parameters (F: smaller negative, and G: Smaller negative) confirm the diagnosis that the bit is balled in a shale interval. The trend of increasingly severe balling based on F and G is shown in Figure 3.4. It is important to notice that drilling performance in this interval is almost the same as for the baseline-1 zone at the beginning of the bit run. ROP, WOB, and Torque have almost the same values in these two sections. As a consequence after concluding that the bit was balled in the interval 12,300'-12,350', it can be determined that the bit was also balled from the beginning of the bit run, and it remained balled until the first section of sand was drilled, at about 12,275', where the bit cleaned up.

Table 3.6 Interpretation of diagnostic parameters for interval 12'300'-12,350'

Data measured at	Conventional Diagnostic parameters			New Diagnostic parameter		Interpretation	Is the Interpretation correct?
	Torque/WOB	Torque/ROP	ROP/WOB	F (Torque, WOB)	G (ROP, WOB)		
Down-hole	Lower	Larger	Lower	Smaller Negative	Smaller Negative	**Bit Balled**	**YES**
Surface	Lower	Larger	Lower	Smaller Negative	Smaller Negative	**Bit Balled**	**YES**

3.1.5. Conclusions and Observations when Evaluating the Method using Down-hole Data

Table 3.7 summarizes the final diagnosis for each of the six sections in the overall interval, 12,250'-12,350', where Aghassi's method was applied. The results from down-hole data and surface data are compared in this table.

The result from this analysis indicated that in four of the six sections, the diagnosis was correct from both down-hole and surface data. The results did not match the interpretation from logs in the interval 12,270'-12'275', where the diagnosis was not clear from either down-hole or surface data. In addition, only the "conventional" diagnostic parameters agreed with log interpretation when using down-hole data in section 12'275'-12,278' and surface data in 12,287'-12,300'.

From Table 3.7 it is possible to determine how often the results from surface data agreed with down-hole data interpretation. In four of the six intervals the lithology interpretation using surface data was the same as for down-hole data. In one of these intervals where they agree (12,270-12,275), the diagnosis does not match log interpretation. They did not fully agree in the sand intervals (12,275'-12,278' and 12,287'-12,300'), where the new diagnostic parameters, F and G, were not conclusive when calculating with down-hole data (interval 12,275'-12,278') and surface data (interval 12'287'-12,300').

Table 3.7 Summary of final diagnosis from Aghassi's method when using down-hole and surface data in Matagorda Island Well # 6, interval 12,250'-12,350'.

Interval (Logging interpretation)	DOWN-HOLE DATA		SURFACE DATA	
	Interpretation	Is correct?	Interpretation	Is correct?
12,250'-12,260' (Shale)	Shale Baseline-1		Shale Base-line-1	
12,260'-12,270' (Shale)	Same as shale Baseline-1	YES	Same as shale Baseline-1	YES
12,270'-12,275' (Shaly Sand)	No clear Diagnosis	NO	No clear Diagnosis	NO
12,275'-12,278' (Sand)	Not clear from New Param.	Only from conventional Par.	Sand	YES
12,278'-12,283' (Shaly Sand)	Sand or clean bit in shale	YES	Sand or clean bit in shale	YES
12,283'-12,287' (Shale)	Shale Baseline-2		Shale Baseline-2	
12,287'-12,300' (Sand)	Sand	YES	Not clear from New Param.	Only from conventional Par.
12,300'-12,350' (Shale)	Balling	YES	Balling	YES

It is important to notice that although the conventional diagnostic parameters (Torque/WOB, Torque/ROP, and ROP/WOB) calculated with down-hole data and surface data do not have the same magnitude, they do have the same value relative to the baseline, see Figure 3.3. Therefore, the final diagnoses for both were always similar. The diagnosis with the new diagnostic parameters, F and G, was not as consistent between surface and down-hole data.

The increase of balling severity is observable in the plots of the new diagnostic parameters F and G versus depth (Figure 3.4) when using down-hole and surface data after the second baseline, but it is not conclusive for the first baseline.

Finally, after evaluating the method with down-hole and surface data, I concluded that acceptable results are obtained when applying Aghassi and Smith's method for diagnosing bit performance using data measured at the surface.

3.2. Application of the Method Using Drilling Data from Wells with Strong Rock Intervals

One of the most important objectives of Aghassi and Smith's method is to provide a means to distinguish between bit balling and strong rock as causes of poor bit performance. Aghassi showed the effectiveness of his method with experimental laboratory data and one field example. I applied the method to four wells from the Gulf of Mexico, Oklahoma, and South America. Each well had a well-known lithology with the presence of some strong rock. After analyzing the results from these wells, two important observations about the method can be made. First, when drilling strong rock, the interpretation of the diagnostic parameters is sometimes that the bit is balled. Second, the results are very sensitive to the selection of the baseline zone. These situations can be observed in the case of the Matagorda Island #1 well, bit run # 15 from 13,200' to 14,500', where shale, bit balling, sand, and strong siltstone are present. This example will be described herein. The objective in this case is to evaluate the responsiveness and consistency of the diagnostic parameters for distinguishing a “balled bit” from “strong rock”.

3.2.1 Definition of Strong Rock

Although there is no standard definition for strong rock, the value of ultimate failure stress in psi and the acoustic travel time in microseconds per foot (usec/ft) from the sonic log can be used to set some values, or ranges of values, for strong rock, especially within a given rock sequence or geologic setting. The greater the ultimates stress in psi and the lower the acoustic travel time in usec/ft, the stronger the rock. In addition, for any single lithology, the lower the porosity the stronger the rock.

The well data used in this research contains a large variation of ultimate stress ranging from 5,000 psi to 53,000 psi depending on lithology, rock properties, formation depth, and the location of the well. Therefore, definitions for strong rock were selected separately for different geologic settings.

A well from Oklahoma has formations with values of ultimate stress from 5,000 psi to 45,000 psi even at shallow depths above 3,000 ft. After correlating the distribution of ultimate stress with drilling parameters and wire-line logs (GR-sonic), I concluded that intervals of strong rock, for this specific well in Oklahoma, can be considered to be the zones with values of ultimate stress greater than 20,000 psi.

One well from South America shows, at about 6,500 ft of depth, interbedded shale (ultimate stress: 8-10,000 psi) and strong sandstone (ultimate stress: 55,000 psi), which is an appropriate environment to apply and evaluate the Aghassi’s method. For that well at that depth, I defined as strong rock the intervals or zones with ultimate stress greater than 30,000 psi. This same criterion was applied to an additional well in the same field in South America.

The data from Matagorda Island # 1 well for the bit run from 13,200' to 14,500' does not have values of formation ultimate stress, However, the gamma ray, LDT-CNL, and sonic logs can be used to make a geological interpretation in order to define the presence of strong rock. This well data was part of the study developed by Smith [9], who made a rigorous diagnosis and description of this bit run. From his analysis and logging interpretation, it can be determined that the siltstones in the intervals 13,315'-13,330' and 13,635'-13,645' are the strongest formations drilled in this bit run. Smith says, "This zone (siltstone) has a very low acoustic travel time of 72 microseconds per foot, equivalent to a compressional velocity of 4.23 kilometers per second which is indicative of a strong rock." Consequently, these two siltstones will be considered as strong rock when applying and evaluating Aghassi's method. Figure 3.5 shows the geological interpretation and drilling data for this interval.

In summary, the definition of strong rock is relative and depends on many factors such as the properties of adjacent rocks, formation depth, and location of the well. However, I expect that the method developed by Aghassi effectively takes into account these factors when the baseline of the bit run is located or updated, so the interpretation of its diagnostic parameters can help to detect strong rock no matter the depth and properties of the formation and the location of the well.

3.2.2 Application of Aghassi's Method to Matagorda Island Well # 1

3.2.2.1 Operational Overview

Aghassi's method was applied using surface data from the Matagorda Island well #1, bit run # 15, for the interval 13,200'-14,500'. Figure 3.5 shows the data from this interval versus depth. The data includes: Rate of penetration (ROP), Torque, weight-on-bit (WOB), rotary speed (RPM), and gamma ray, sonic, neutron NPHI and density DPHI logs.

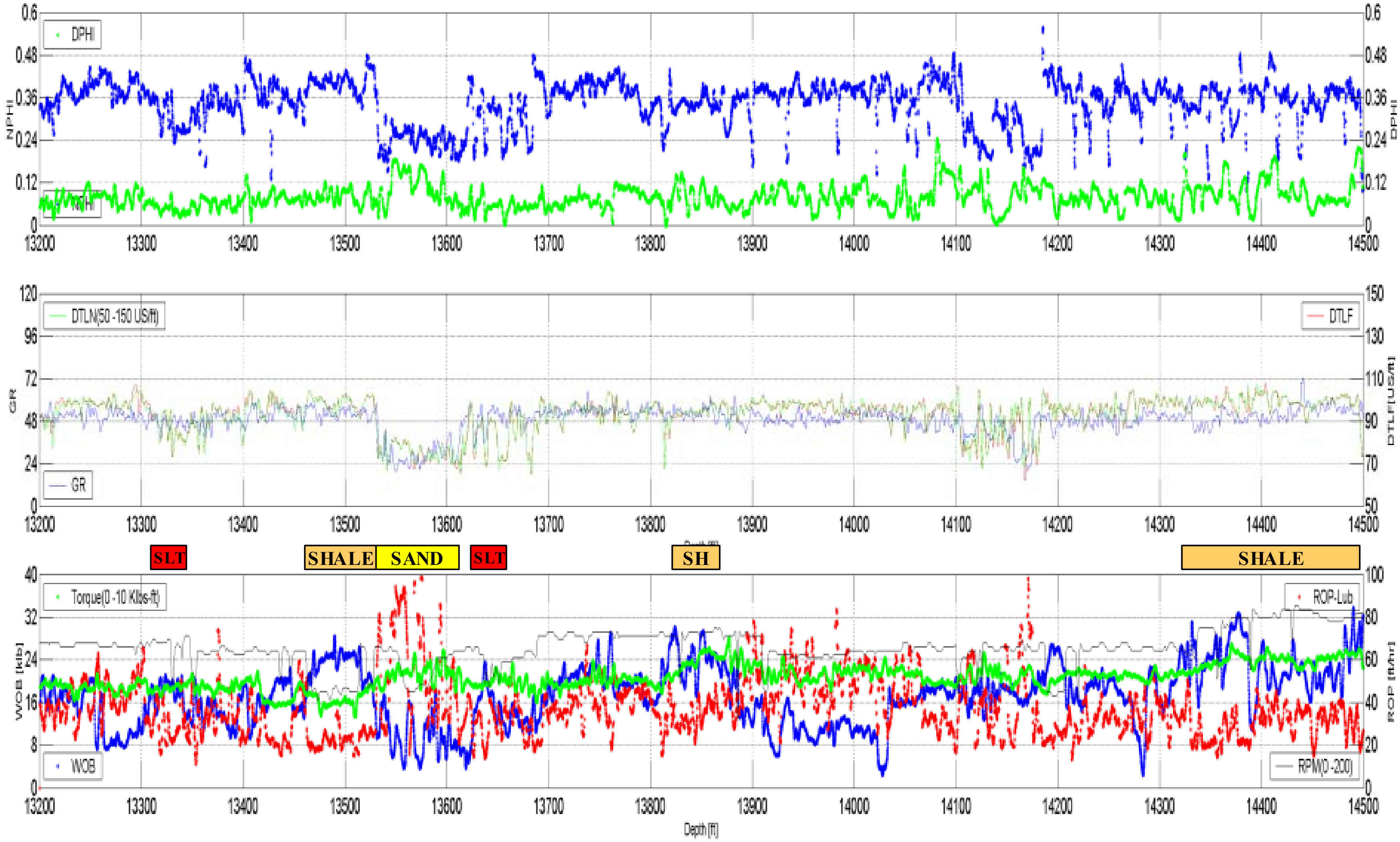

Figure 3.5 Drilling Data and Logs from Matagorda Island Well #1.

The interval was drilled using a 12 1/4" HCC-AR554 bit with three 12/32" and three 13/32" jets. After drilling 1300', the bit was evaluated using the dull grading system as 3-5-RO-T-X-I-BT-PR. The bottom-hole-assembly had five 9 ½" drill collars, fourteen 8" drill collars, a drilling jar, and nine joints of 5" Hevi-Wate drill pipe. The average operating parameters were 120-140 rpm 750-gpm flow rate and 3500 psi of pump pressure. The drilling fluid was water-based mud with a density of 16.2 ppg.

According to the interpretation of gamma ray and porosity logs, this interval is mainly shale with two sandstones, at 13,522'-13,600' and 14,155'-14,175', which were drilled at the highest penetration rate of 65 ft/hr average. The WOB during this bit run varies from low values of 5,000 lbs, used to drill the sandstones, to high values up to 32,000 lbs used to drill shale with poor bit performance. The siltstone sections, denominated as SLT in Figure 3.5, were drilled with an average WOB of 20,000 lbs.

3.2.2.2 Sequential Summary of Drilling Parameters and Lithology

Taking into account the lithology and the behavior of the drilling parameters shown in Figure 3.5, the following descriptive summary of the bit run can be made. The bit run started at 13,200' with a average WOB of 16,000 lbs in a shale section with a good average penetration rate of 40 ft/hr during the first 100 feet. Once the bit entered the first siltstone interval (13,315'-13,330'), the ROP decreased to a value of 20 ft/hr, and the WOB had to be increased to 20,000 lbs. The next long shale interval (13,330'-13,460') was drilled with an average ROP of 30 ft/hr and WOB of 15,000 lbs. The subsequent interval (13,460'-13,522') is also shale, but there the bit performance was different. At 13,460', the ROP decreased to 20 ft/hr, and the WOB was increased gradually to 24,000 lbs trying to improve bit performance, but the ROP kept its low value of 20 ft/hr. In addition, torque decreased from 6,000 ft-lbs to 4,000 ft-lbs in this interval.

All of these signals are symptoms that the bit is probably becoming balled in this section. Once the bit entered the first sand (13,522'-13,600') and began to clean up, the ROP began to increase significantly reaching a maximum value of 90 ft/hr at 13,550' with a low WOB of 5,000 lbs. After that, the ROP decreased and at the end of the sand section had an average value of 45 ft/hr. Then a short section (13,600'-13,635') of interbedded shale and sand was drilled with ROP of 30 ft/hr and WOB of 8,000 lbs. Next, the second siltstone (13,635'-13,645') was drilled, where the ROP averaged 30 ft/hr, but the WOB was increased to 24,000 lbs apparently because of the stronger rock.

From this point to the end (14,500') of the bit run, there are additional sections of shale and sand. The performance shows ROP varying from 20 ft/hr in shale to 60 ft/hr in clean sand. Apparently, the poor bit performance in shale sections occurred because the bit was at least partially balled. However, variations in bit performance in the shale indicate that this condition was not as pervasive as in the bit run analyzed in section 3.1. The last 100 feet (14,400'-14,500') of this bit run is a shale interval which was drilled with a high average WOB of 30,000 lbs, and the ROP decreased to an average of 20 ft/hr. This decrease could occur because the bit was balled or because the bit was considerably worn at that time.

3.2.2.3 Selection of the Baseline

The evaluation of this bit run was made setting the baseline in the interva1 13,265'-13,295' which is a shale drilled with relatively high ROP of about 40 ft/hr and a WOB of 8,000 lbs. Because of these good drilling conditions, the entire bit run was evaluated considering only one baseline. In other words, there is a relatively long shale drilled with good bit performance at the beginning of the bit run. These are the favorable conditions described by Aghassi[1] for locating the baseline. In addition, there are no drastic changes in drilling parameters suggesting that the baseline has to be updated during the bit run.

3.2.2.4 Calculation and Evaluation of Diagnostic Parameters

The data from Matagorda Island # 1 Well was only available on logs plotted versus depth. Consequently, the method was applied and validated with this data by converting the depth-based data to time-based data. Plots were digitized at 0.1 ft intervals and depths were converted to times dividing by the average value of ROP. The derivative parameters (F, G, and dWOB/dt) were calculated using linear regression with the 20 previous points.

Figure 3.6 shows plots of the drilling parameters (WOB, torque, RPM, and corrected ROP), the conventional diagnostic parameters (Torque/WOB, Torque/ROP, ROP/WOB), and the new diagnostic parameters (F and G) versus depth for the entire bit run. The interval selected as a baseline is shown on the plots of conventional and new diagnostic parameters. In addition, the lithology interpretation of the six intervals where the method was evaluated is also shown. The siltstone intervals are denominated as SLT in this figure. Table 3.8 summarizes the interpretation of the diagnostic parameters for each interval studied for this bit run.

Once the baseline-1 (13,265’-13,295’) is defined, the diagnostic parameters are calculated and evaluated for each of the six intervals shown in figure 3.6.

- 13,310’-13,330’: According to the geological interpretation, this is the first interval of strong siltstone. However, the diagnosis from the conventional parameters is that the bit is balled. Torque/WOB and ROP/WOB are lower, while Torque/ROP is larger, than baseline values, see Figure 3.6 and Table 3.8.

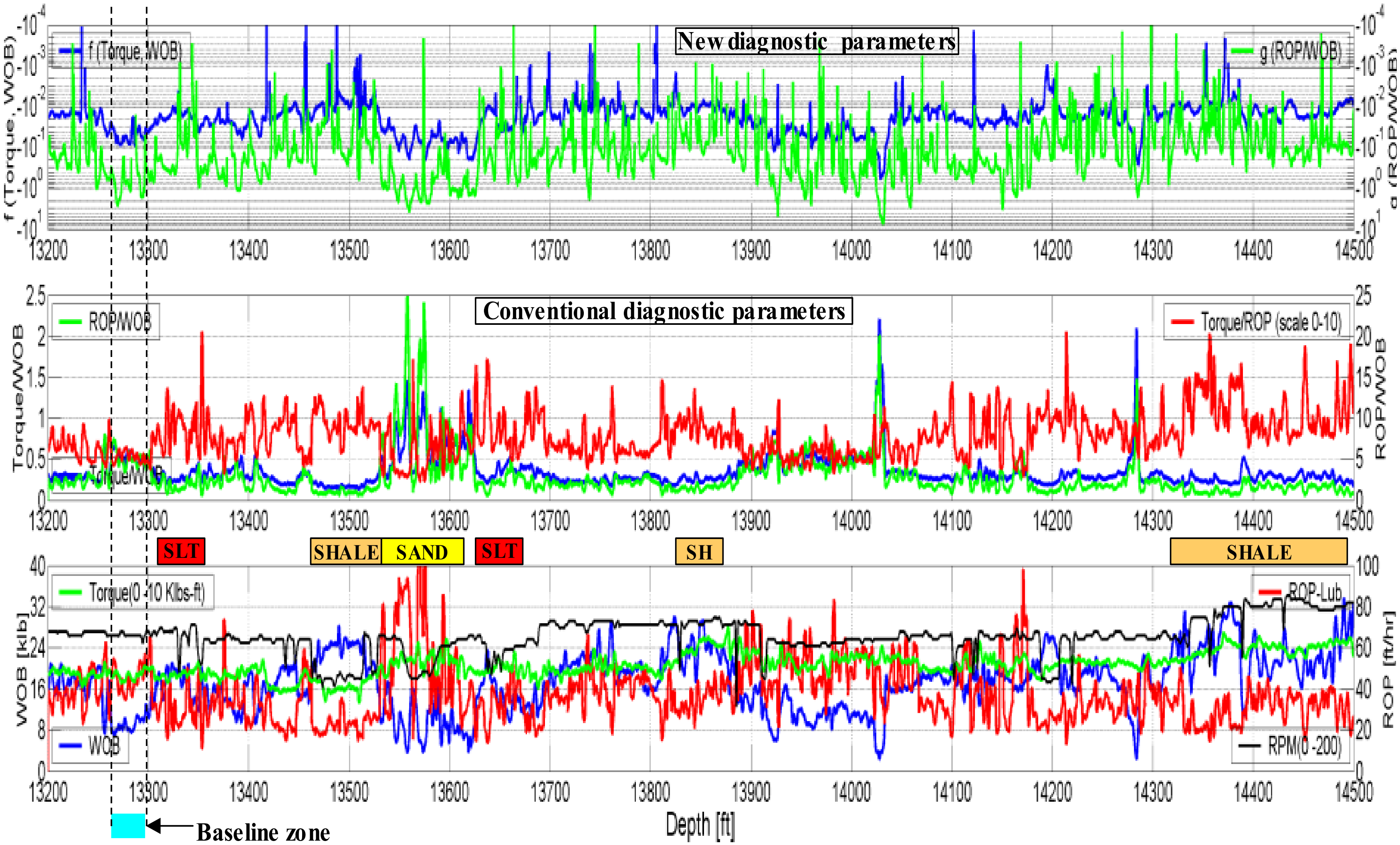

Figure 3.6 Conventional and New Diagnostic Parameters for Matagorda Island Well #1. Baseline Located at 13,265'-13,295.

Table 3.8. Incorrect Diagnosis of Strong Siltstone as Bit balled. Baseline located at 13,265'-13,395'.							
Interval (Logging interpretation)	**Conventional Diagnostic Parameters**			**New Diagnostic Parameters**		**Interpretation**	**Is it correct?**
	Torque/WOB	**Torque/ROP**	**ROP/WOB**	**f (Torque, WOB)**	**g (ROP, WOB)**		
13'265'-13,295' (Shale)	**Baseline-1**	**Baseline-1**	**Baseline-1**	**Baseline-1**	**Baseline-1**	**Shale Base-line-1**	
13,310'-13,330' (Silstone)	Smaller	Larger	Smaller	Smaller negative	Smaller negative	**BALLING**	NO
13,460'-13,522' (Shale)	Smaller	Larger	Smaller	Smaller negative	Smaller negative	**BALLING**	**YES**
13,522'-13,600' (Sand)	Larger	Same-Larger	Larger	Larger negative	Larger negative	**SAND**	**YES**
13,613'-13,645' (Silstone)	Smaller	Larger	Smaller	Smaller negative	Smaller negative	**BALLING**	NO
13'830'-13,860' (Shale)	Smaller	Larger	Smaller	Smaller negative	Smaller negative	**BALLING**	**YES**
14,320'-14,500' (Shale)	Smaller	Larger	Smaller	Smaller negative	Smaller negative	**BALLING**	**YES**

In addition, the response of the new diagnostic parameters, F and G, support the diagnosis of a balled bit for this interval, because they are smaller negatives than baseline values. This result does not agree with the geological interpretation that this interval is strong rock.

This incorrect diagnosis using Aghassi's method could occur for several reasons. First of all, consider the basis for diagnosis given in Table 2.1. Three of the five diagnostic parameters (Torque/ROP, ROP/WOB, and G) are similar for strong rock and bit balling situations. Then, only two diagnostic parameters are useful for distinguishing between strong rock and bit balling. They are the conventional parameter "Torque/WOB", which is smaller for balling and larger for strong rock, and the new parameter "F", which is a smaller negative for balling and a larger negative for strong rock. These two parameters are both functions of Torque. Then, a change in the response of Torque measured at surface, when the bit passes from shale or weak rock to strong rock, is required to get a good diagnosis from these parameters. However as seen in Figure 3.6, the Torque has an almost constant value of 5,000 ft-lb during the first 200 ft of this bit run where the first interval of strong siltstone is located.

As a consequence, this lack of response of surface torque when the lithology change occurred affects the results of the final diagnosis of the method. Another factor that could have affected the diagnosis is the selection of the baseline zone. This factor will be studied in the next section of this chapter.

- 13,460'-13,522': As described previously, the ROP in this shale interval decreased to a value of 20 ft/hr while the WOB is increased gradually up to 24,000 lbs, and torque decreased from 6,000 ft-lb to 4,000 ft-lb. Then, the behaviors of these drilling parameters are symptoms of poor bit performance, and maybe the bit is getting balled. The results of Aghassi's method corroborated this situation. The diagnosis from the conventional diagnostic parameters is that the bit is balled, because Torque/WOB and ROP/WOB are smaller while Torque/ROP is larger than their values in the baseline, see Table 3.8. In addition, the new diagnostic parameters ratified the bit-balling situation because "F" and "G" have a smaller negative value than their value in the baseline zone.

- 13,522'-13,600': This interval of clean sand is clearly detected by the GR- neutron NPHI-density DPHI logs, see Figure 3.5. As soon the bit entered this sand, the ROP increased from that in the previous shale where the bit was balled. In addition, in this sand interval the bit had its highest ROP (90 ft/hr) with the lowest WOB (5,000 lbs). As reported in the Table 3.8, the diagnosis from the method is correct because the responses of the five diagnostic parameters observed in Figure 3.5 are typical symptoms of bit performance in clean sand.

- 13,635'-13,645': According to the geological interpretation, this is the second interval of siltstone or strong rock during the bit run. As for the previous siltstone interval (13,310'-13,330') the diagnosis from the method is that the bit is balled. As reported in the Table 3.8, this diagnosis is not correct. The potential causes of this lack of accuracy seem to be the same as for the previous siltstone.

- 13,830'-13,860': A long shale zone is encountered from approximately 13,650' to 14,100'. The interval (13,830'-13,860') was selected as typical of this zone, and Smith[9] also described it as shale. From the analysis of the conventional diagnostic parameters, it can be observed that Torque/WOB and ROP/WOB have lower values than their baseline values. In contrast, Torque/ROP is larger; see Figure 3.6. In addition, the new diagnostic parameters "F" and "G" have a smaller negative value than the baseline. These are all symptoms of a balled bit. Although this interval was drilled with an average ROP of 30 ft/hr, and the Torque response shows a higher value than its value in the baseline zone, the final diagnosis of the method is that the bit is balled. Although this ROP is greater than 25 ft/hr, it is relatively low and the WOB relatively high with in the bit run, therefore the diagnosis is concluded to be correct.

- 14,320'-14,500': This shale interval is the final zone of this bit run, and it was drilled with poor bit performance. The average ROP and WOB when drilling this interval were 20 ft/hr and 30,000 lbs respectively. In addition, the RPM were increased from 120 to 170, apparently trying to get a better ROP with no success. As shown in Table 3.8, the diagnosis from all five of the diagnostic parameters of Aghassi's method is that the bit is balled. Given that the bit was "ringed out" when pulled, the actual cause for the decrease in performance could be bit balling, bit wear, or both.

After the evaluation of these intervals from Matagorda Island Well # 1, it can be concluded that the method supplied good results for all the situations except for strong rock. The interpretation was not correct when drilling strong rock, where the diagnosis was that the bit was balled. This result was also obtained when the method was evaluated for drilling in strong rock in two additional wells. One of the possible causes of this erratic diagnosis can be the location of the baseline. In other words, the final diagnosis from the method can be affected by the selection of the baseline. For that reason, I studied the effect of changing the location of the baseline to another zone.

3.2.2.5 The Effect of Selecting the Baseline Location

In order to evaluate the effect of baseline selection on the final result, the same data from the bit run # 15 of Matagorda Island Well # 1 was considered, but the baseline was located at the very beginning of the bit run, in the interval 13,210'-13,240', where there is a shale with relatively high ROP (40 ft/hr average).

Figure 3.7 repeats the plots for drilling, conventional, and new diagnostic parameters from Figure 3.6 with the new baseline location. As seen in figure 3.7, for the two siltstone intervals (13,310'-13,330' and 13,613'-13,645) the value of the conventional parameter Torque/WOB is very close to its new baseline value. In addition, Torque/ROP is still larger and ROP/WOB is now slightly smaller. Then, the diagnosis from the conventional parameters tends to be that these intervals are strong rock. The new diagnostic parameter F ratified this diagnosis because its value is the same as the baseline value. Table 3.9 summarized the interpretation of the five diagnostic parameters for all six intervals. Now, the diagnosis in the two strong intervals is correct, as well as the diagnosis of the other intervals.

Consequently, the diagnosis was improved due to the new location of the baseline. Therefore, the accuracy of the method is sometimes dependent on the selection of the baseline.

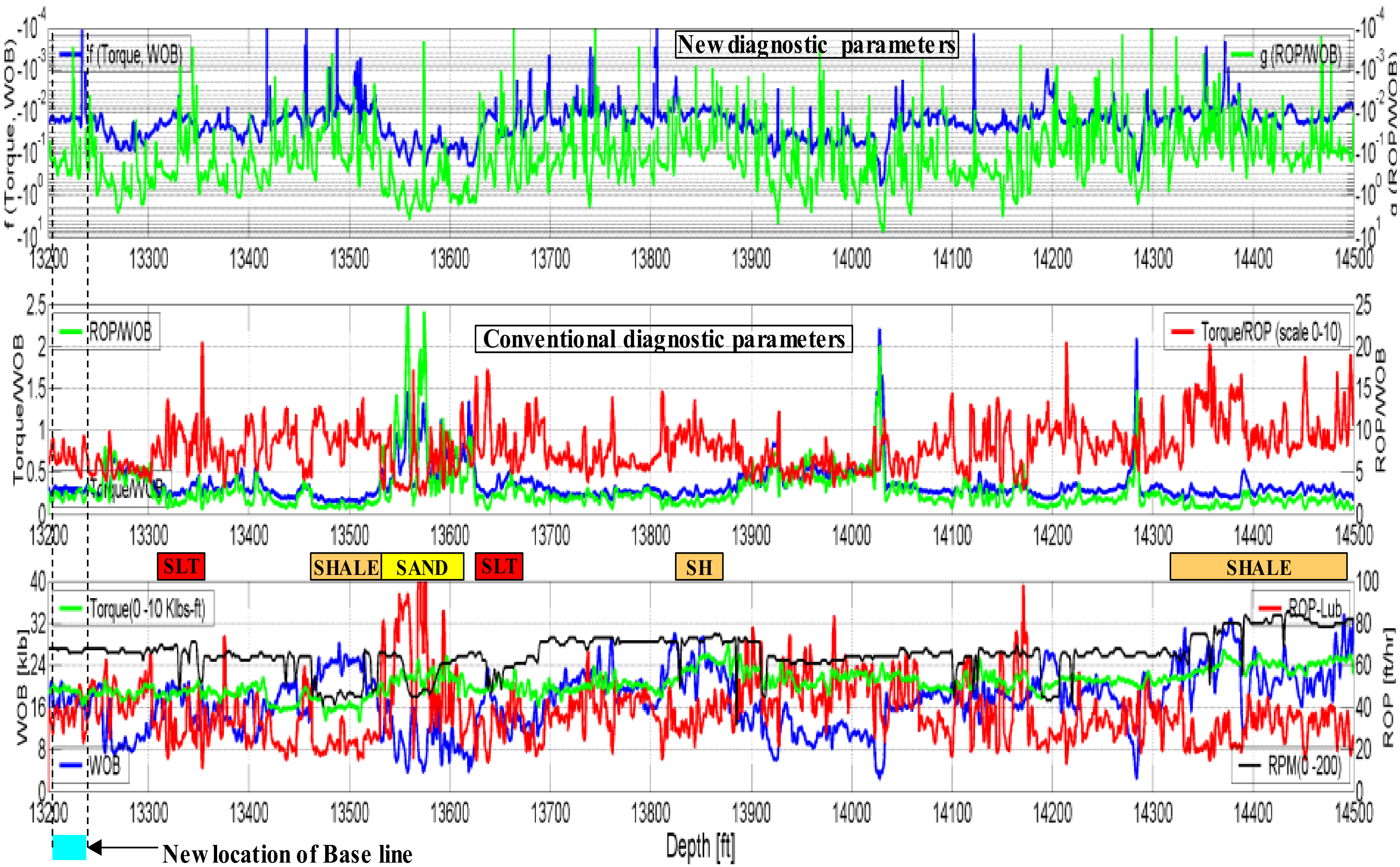

Figure 3.7 Correct Diagnosis for Matagorda Island Well #1 with baseline located at 13,210'-13,240'.

Table 3.9. Correct Diagnosis of Strong Siltstone as Bit balled. Baseline located at 13,210'-13,240'.								
Interval (Logging interpretation)	Conventional Diagnostic Parameters			New Diagnostic Parameters		Interpretation	Is it correct?
	Torque/WOB	Torque/ROP	ROP/WOB	f (Torque, WOB)	g(ROP, WOB)		
13'210'-13,240' (Shale)	Baseline-1	Baseline-1	Baseline-1	Baseline-1	Baseline-1	Shale Base-line-1	
13,310'-13,330' (Silstone)	Close	Larger	Slightly Smaller	Same	Slightly Smaller negative	STRONG ROCK	YES
13,460'-13,522' (Shale)	Smaller	Larger	Smaller	Smaller negative	Smaller negative	BALLING	YES
13,522'-13,600' (Sand)	Larger	Same-Larger	Larger	Larger negative	Larger negative	SAND	YES
13,613'-13,645' (Silstone)	Close	Larger	Smaller	Same	Smaller negative	STRONG ROCK	YES
13'830'-13,860' (Shale)	Smaller	Larger	Smaller	Smaller negative	Smaller negative	BALLING	YES
14,320'-14,500' (Shale)	Smaller	Larger	Smaller	Smaller negative	Smaller negative	BALLING	YES

There is an important difference in one of the drilling parameters between the two baseline zones. In the first baseline selected (13,265-13,295), a low average WOB of 8,000 lbs was used. On the other hand, the second selection (13,210-13,240) has a relative high average WOB of 18,000 lbs. It is hypothesized that the value of the WOB is critical in the selection of the baseline because it affects the diagnosis from the parameters "Torque/WOB" and "F" which are the most important for distinguishing between bit balled and strong rock.

Apparently, if the baseline is set in an interval with relatively low WOB, such as 13,265-13,295, the Torque/WOB in zones of stronger rock is likely to be smaller than the baseline value. This is because although WOB and Torque usually increase in stronger formations, the proportional change in WOB is almost always greater than the change in Torque. Consequently, if the increase of WOB is greater than Torque increase, Torque/WOB tends to become smaller, and the diagnosis is a balled bit instead of strong rock, as in Table 3.8. On the other hand, if the baseline is in a section drilled with relatively high WOB, like 13,210-13,240, the change in WOB when drilling strong rock is smaller. Then, Torque/WOB is likely to be slightly larger or close to its baseline value, as in Table 3.9.

In summary, it is hypothesized that a better interpretation of the strong rock sections can be obtained if the baseline is located in a shale section with both a relatively high ROP, as Aghassi recommended, and a relatively high WOB.

3.2.3 Conclusions and Observations when Evaluating the Method for Strong Rock

The method was also applied to three more wells with known lithology and presence of strong rock, two from South America and one from Oklahoma,. The information available for these wells includes drilling data, GR-sonic-SP-resistivity logs, samples analysis, and geological description. Summaries of the final results for these wells are shown in Table 3.10, Table 3.11, and Appendix I.

Table 3.10 illustrates the results obtained for the L#12D Well located in South America. The 8 ½" bit run analyzed drilled three sections of strong chert and two shale sections where the bit was getting balled. The method was applied in these five intervals as well as in an interval interpreted as shale clean bit. As seen in the table, two of the three strong rock intervals were diagnosed as balled bit. For the remaining four intervals the diagnosis was correct.

Table 3.10. Diagnosis from Aghassi's Method for L # 12 D Well , South America.

Interval (Logging interpretation)	Conventional Diagnostic Parameters			New Diagnostic Parameters		Interpretation	Is it correct?
	Torque/WOB	Torque/ROP	ROP/WOB	F (Torque, WOB)	G (ROP, WOB)		
6,322'-6,340 (Shale)	**Baseline-1**	**Baseline-1**	**Baseline-1**	**Baseline-1**	**Baseline-1**		
6,415'-6,435 (Strong chert)	Larger	Larger	Smaller	Larger negative	Noisy	Stronger Rock	YES
6,530'-6,560' Shale getting balled	Slightly Smaller	Larger	Smaller	Smaller negative	Noisy	Getting Balled	YES
6,560'-6,630' (Strong chert)	Smaller	Larger	Smaller	Smaller negative	Smaler Negative	BALLING	NO
6,680'-6,710' (Shale)	**Baseline-2**	**Baseline-2**	**Baseline-2**	**Baseline-2**	**Baseline-2**		
6,740'-6,800' (Shale clean bit)	Larger	Same	Larger	Larger negative	Larger negative	Shale Clean Bit	YES
6,825'-6,840' Shale getting balled	Smaller	Larger	Smaller	Slightly Smaller Neg.	Smaler Negative	BALLING	YES
7,300'-7,440' (Strong chert)	Smaller	Larger	Smaller	Smaller negative	Smaler Negative	BALLING	NO

Table 3.11 shows the diagnosis for the intervals in L # 5 A Well, located in the same field in South America. The 8 ½" bit run drilled two sections of strong chert and one shale section where the bit was getting balled. The method was applied in these three intervals as well as in three intervals interpreted as shale drilled with a clean bit. As seen in the table, one of the two strong rock intervals was diagnosed as balled bit. In addition, one interval of clean shale was diagnosed as strong rock. For the remaining four intervals, the diagnosis was correct.

Table 3.11 Diagnosis from Aghassi's Method for L # 5 A Well , South America.

Interval (Logging interpretation)	Conventional Diagnostic Parameters			New Diagnostic Parameters		Interpretation	Is it correct?
	Torque/WOB	Torque/ROP	ROP/WOB	F (Torque, WOB)	G (ROP, WOB)		
6,450'-6,490' (Shale)	**Baseline-1**	**Baseline-1**	**Baseline-1**	**Baseline-1**	**Baseline-1**		
6,550'-6,600 (Strong chert)	Larger	Larger	Close	Larger negative	Larger negative	Stronger Rock	**YES**
6,650'-6,740' (Shale)	Close	Larger	Smaller	Larger negative	Noisy	Stronger Rock	**NO**
6,740'-6,770'. (Shale)	**Baseline-2**	**Baseline-2**	**Baseline-2**	**Baseline-2**	**Baseline-2**		
6,880'-6,920' (Shale clean bit)	Larger	Same	Smaller	Slightly Larger neg.	Noisy	Shale Clean Bit	**YES**
6,950'-7,130' (Shale clean bit)	Larger	Close	Smaller	Larger negative		Shale Clean Bit	**YES**
7,200'-7,250' Shale getting balled	Smaller	Larger	Smaller	Smaller negative		BALLING	**YES**
7,270'-7,400' (Strong chert)	Smaller	Larger	Smaller	Smaller negative	Smaler Negative	BALLING	**NO**

The diagnosis for the Oklahoma Well is shown in Appendix I. The method was applied in four bit runs including seven intervals of stronger rock. In two of these intervals, the diagnosis was not conclusive. In the five remaining intervals, the diagnosis was correct.

After evaluating several bit runs of different wells from different areas where intervals of strong rock where drilled, the following conclusions and observations are formulated:

- Often, when drilling strong rock, the interpretation of the diagnostic parameters is not correct because the diagnosis is that the bit is balled.

- One of the apparent situations that causes this error is a lack of variation of the torque measured at surface when changing from a weak formation to strong rock. Apparently, that happens in the field when the driller concluded that a strong rock was encountered and decreased RPM and increased WOB as the ROP decreased. If torque does not increase as expected in strong rock, the interpretation is that bit balling is occurring. Therefore, the response of the two parameters, "Torque/WOB" and "F" which distinguish between strong rock and bit balling, is different than expected.

- The accuracy of the diagnosis when applying Aghassi and Smith's method is sometimes dependent on the selection of the baseline zone. The value of the WOB is apparently important in the selection of the baseline, because it affects the diagnosis of the parameters "Torque/WOB" and "F". If the baseline is set in an interval with relatively low WOB, the Torque/WOB in zones of stronger rock will be smaller or lower than its baseline value because the increase in WOB will be proportionately higher than the increase in Torque when drilling the stronger rock. The diagnosis is then a balled bit instead of strong rock. On the other hand, if the baseline is in a section drilled with relatively high WOB, the Torque/WOB is likely to be slightly larger or close to its value in the baseline, as required for a correct diagnosis of strong rock. Then, besides Aghassi's advice about locating the baseline in a shale interval with relatively high ROP, I recommend setting the baseline, preferably, over an interval of relatively high WOB.

3.3 Reliability of Aghassi's Method

One way to evaluate the reliability of Aghassi's method is calculating the percent of intervals correctly diagnosed in wells with known lithology. For example in Matagorda Island Well # 6, four of the six intervals analyzed were correctly diagnosed**.** The interval 12,270'-

12,275' has no clear diagnosis from the method. In addition, one sand section has incorrect diagnosis from the new parameters. Then, the reliability of the method in this well was 4/6 or 67%.

In Matagorda Island Well # 1, four of the six intervals analyzed were correctly diagnosed, because two sections of strong rock were interpreted as a balled bit. Then the reliability of the method was also 4/6 or 67% for this well.

As mentioned previously, the method was also applied in a well from Oklahoma and two from South America. The well from Oklahoma will be described in the next chapters. As seen in Appendix I, in this well 21 intervals were analyzed using Aghassi's method, 7 were correctly diagnosed and 3 were wrong. The 11 remaining intervals were partially diagnosed, because they were correctly diagnosed only for one of the set of diagnostic parameters, conventional or new, then the diagnosis for these intervals was not conclusive. Consequently, the efficiency for this well was 7/21 or 33%.

The L # 12A and L # 5D wells from South America were also analyzed. In the L # 12A well, four of six intervals were correctly diagnosed, because two interval of strong rock were diagnosed as bit balled, see Table 3.10. The efficiency of the method in this well was 4/6 or 67%. In the L # 5D well, the efficiency was also 4/6 or 67 % because for two of six intervals the diagnosis was not correct, see Table 3.11.

Table 3.12 summarizes the evaluation of the results obtained after appying Aghassi's method in 5 wells. The reliability of the method was calculated not only for each well but also for each of the situations or lithology diagnosed by the method: sand, shale clean bit, shale bit balling, and strong rock. As seen in this table, the highest reliability (82%) of the method was obtained for bit balling prediction, because 9 of the 11 intervals where the bit was really balled were diagnosed correctly.

The reliability of method for prediction of drilling sand and shale with a clean bit was 50% and 40% respectively. 11 strong rock intervals where analyzed, and only 4 of them had correct prediction. Consequently, the method has the lowest reliability (36 %) for strong rock diagnosis.

In summary, 45 intervals were analyzed with Aghassi's method in the five wells, about 9 intervals per well. 23 of these intervals were correctly diagnosed. Consequently, the overall efficiency of Aghassi's method was 23/45 or 51%, and the most common cause of error was the wrong diagnosis of strong rock and shale drilled with a clean bit.

In response to these shortcomings of Aghassi's method, I studied and applied a new technique for diagnosing bit performance. Statistical, binary logistic regression models were evaluated for distinguishing between strong rock and a balled bit. The methodology and the results obtained from these statistical models are explained in the next chapters.

Table 3.12 - Summary and Evaluation of Aghassi's Method													
WELL / Situation	Matagorda I # 6		Matagorda I # 1		Oklahoma		L # 12 D		L # 5 A		TOTAL		Reliability per Situation
	No. of known occurences	No. Diagnosed correctly	No. of known occurences	No. Diagnosed correctly	No. of known occurences	No. Diagnosed correctly	No. of known occurences	No. Diagnosed correctly	No. of known occurences	No. Diagnosed correctly	No. of known occurences	No. Diagnosed correctly	
Sand	3	2	1	1	4	1					8	4	**50%**
Shale (Cleaner Bit)	2	1			9	2	1	1	3	2	15	6	**40%**
Shale (Severe Balling)	1	1	3	3	5	3	1	1	1	1	11	9	**82%**
Stronger rock			2	0	3	1	4	2	2	1	11	4	**36%**
TOTAL	6	4	6	4	21	7	6	4	6	4	45	23	**51%**
Reliability per Well	**67%**		**67%**		**33%**		**67%**		**67%**		**51%**		

4. THE LOGISTIC REGRESSION

4.1. Introduction

Statistical logistic regression models were studied and applied in this research as an alternative way to diagnose bit performance. This kind of regression was developed to overcome problems encountered when applying the well-known linear regression model[33, 34]. Nowadays, the logistic regression is a powerful device for use in cases where the linear regression provides inefficient estimation.

In this chapter, the logistic regression is introduced starting with the principles of linear regression including assumptions, calculation of coefficients, and model evaluation. The failings of linear regression when the dependent variable of the model is dichotomous, divided into two mutually exclusive groups or options, are explained. Then, the logistic model equation is derived. In addition, the calculation and interpretation of the logistic model's coefficients as well as the evaluation of the model are described. A parallel between linear and logistic regression is made. It allows a better understanding of this new topic, by stepping from a familiar terrain, the linear regression, to a new turf, the logistic regression.

4.2. Regression

The purpose of regression is to produce a model to predict the value of a property from the measurements of other related properties. The regression procedure consists of the following steps:

1) Generation of data or measurements of both the observed quantity and the predicted quantity.

2) Derivation or assumption of a model, which is an equation that relates one or more observed quantities to a quantity to be predicted. The model contains unknown parameters or coefficients that must be evaluated.

3) Selection of a method to determine the unknown parameters or coefficients of the model, which are calculated from the measurements of all the quantities involved.

4) Application of the model to predict the desired quantity.

Examples of this kind of model are the linear regression model and the logistic regression model.

4.3. Linear Regression Model

The most common type of regression is linear regression. In linear regression analysis, it is possible to test whether two variables are linearly related and to calculate the strength of the linear relationship when the relationship between the variables can be described by the following equation:

$$\mathbf{Y} = \boldsymbol{a} + \boldsymbol{B}\,\mathbf{X} \qquad \text{(4.1)}$$

Where **Y** is the variable being predicted or the "dependent variable". **X,** is the variable whose values are being used to predict **Y**, or also called the "independent variable." ***a*** and ***B*** are parameters or coefficients which have to be calculated mathematically. The parameter ***a***, called the intercept, represents the value of **Y** when **X** = 0. The parameter ***B***, is the slope of the line that provides the best linear estimate of **Y** from **X**, and it represents the change in **Y** associated with a one-unit increase in **X.**

When there are several independent variables, the multiple linear regression is represented by the following equation:

$$\mathbf{Y} = \boldsymbol{a} + \boldsymbol{B_1}\,\mathbf{X_1} + \boldsymbol{B_2}\,\mathbf{X_2} + \ldots\ldots \boldsymbol{B_n}\,\mathbf{X_n} \qquad \text{(4.2)}$$

Where **n** is the number of independent variables $\mathbf{X_1}, \mathbf{X_2} \ldots \mathbf{X_n}$.

B_1, B_2……B_n are called the partial slope coefficients because each of the **n** independent variables $X_1, X_2 \ldots X_n$ provides only a partial influence or prediction for the value of **Y**.

Equation (4.1) is sometimes written in a form that explicitly recognizes that that the prediction of **Y** from **X** may be imprecise:

$$Y = a + B\,X + e \quad \ldots\ldots \quad (4.3)$$

For several **n** independent variables, equation (4.2) becomes:

$$Y = a + B_1 X_1 + B_2 X_2 + \ldots\ldots B_n X_n + e \quad (4.4)$$

Where **e** is the error term, a random variable that represents the error in predicting **Y** from **X.**

4.3.1 Calculation of the Linear Regression Coefficients

The method of ordinary least squares (OLS) estimation is used to calculate the intercept a and the regression coefficients B_1, B_2……B_n. The least squares criterion is a minimization of the sum of the squared differences between the observed responses, Y_i , and the predicted responses, $\hat{Y}_i$, for each fixed value of X_i. The differences, $Y_i - \hat{Y}_i$, are called residual e_i. In other words, the OLS produce the equation $\hat{Y} = a + B\,X$, or in the case of several independent variables, $\hat{Y} = a + B_1 X_1 + B_2 X_2 + \ldots B_n X_n$, where $\hat{Y}$ is the value of the **Y** predicted by the linear regression equation.

In mathematical terms, the error e_i for a bivariative regression can be expressed as function of a and B as follow:

$$e\,(a, B) = \sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - a - B\,X_i)^2 \quad \ldots\ldots \quad (4.5)$$

We want to find the values of the slope (B) and the intercept (a) that make the residual **e** as small as possible.

A differential change in **e**, d**e**, can be written in terms of the differential changes in of ***a*** and ***B*** as :

de = (∂e/∂*a*) d*a* + (∂e/∂*B*) d*B* ……………………………………...……(4.6)

The minimal **e** occurs when **de** = 0. After making **de** = 0, the following equations are obtained:

$$\alpha = \frac{1}{k}\sum_{i=1}^{k} Y_i - \beta \frac{1}{k}\sum_{i=1}^{k} X_i \quad \text{……………………………… (4.7.a)}$$

$$\beta = \frac{\sum_{i=}^{k} X_i Y_i - \frac{1}{k}(\sum_{i=1}^{k} X_i)(\sum_{i=1}^{k} Y_i)}{\sum X_i^2 - \frac{1}{k}(\sum_{i=1}^{k} X_i)^2} \quad \text{……………………………..(4.7.b)}$$

α and β are the values of ***a*** and ***B*** that minimize the error **e**, and they also have a statistical interpretation. α and β are the slope and intercept estimated derived from a sample of k data; therefore, they may differ from the population slope and intercept values. In addition, they give the model or equation:

Ŷ = α - β **X** ……………………………………………………….. (4.8)

For multiple regression analysis with **n** independent variables, it is necessary to resolve a multiple system of **(n+1)** simultaneous equations for calculating the **(n+1)** unknown parameters (α, β_1, $\beta_2 \ldots \beta n$). This system is usually solved with the Gauss elimination method which is explained is several basic statistics and numerical methods books [38].

4.3.2 Evaluation of the Linear Regression Model

Linear regression models are evaluated by calculating the Coefficient of Determination (R^2), which indicates how well we can predict the dependent variable from the independent variable using linear regression models[35]. The coefficient of determination ranges in value from 0 to 1. If $R^2 = 1$ or close to one, there is a perfect correlation model-sample data, then the prediction of **Y** from **X** is accurate. At the other extreme, if the coefficient of determination is 0 or close to 0, the regression equation is not helpful in predicting the value of the dependent variable, **Y**.

4.3.3 Linear Regression Assumptions

In order to use the ordinary least square (OLS) method to estimate the coefficients in linear regression, several specific assumptions must be satisfied [34], three of them are the following:

1) Measurement: All independent variables are interval, ratio, or dichotomous; and the dependent variable is continuous, unbounded, and measured on an interval or ratio scale. All variables are measured without error.

2) Normality of errors: The errors are normally distributed for each set of values of the independent variables.

3) Homoscedasticity: The variance of the error term, **e**, is the same or constant for all values of the independent variables.

4.3.4 Dichotomous Variables in Linear Regression

Dichotomous means divided into two mutually exclusive groups or options. The independent and the dependent variables of a model can be dichotomous. According to Menard[34], the linear regression models work well when the independent variable is dichotomous, and the interpretation of results makes substantive as well as statistical sense.

On the other hand, when the dependent variable is dichotomous, the interpretation of the linear regression model is not as straightforward. When there is a dichotomous dependent variable, it is convenient to code the values of the variables as 0 and 1, because it produces the result that the mean of the variable is the proportion of cases with a value of 1, and the predicted value of the dependent variable can be interpreted as the predicted probability. Then, the values of the dependent variable, **Y**, must have values between 0 and 1. But, linear regression models with a dichotomous dependent variable violate the assumptions of "normality" and "homoscedasticity" described previously. Besides, the models provide values of **Y** greater than 1 and lower than 0 (negatives values), which does not make statistical sense.

Consequently, using linear regression models with a dichotomous dependent variable is inappropriate. Then, it is evident that a non-linear model is better suited to the analysis of the dichotomous dependent variable. Also, it is necessary to consider alternative methods or techniques for estimating the parameters or coefficients to describe the relationship between **Y** (dichotomous dependent variable) and **X** (independent variable).

4.3.5 Solutions and an Alternative Model for Dichotomous Dependent variables

Linear regression provides unsuitable results for models with a dichotomous dependent variable. Some of the problems generated are related with boundaries (values of **Y** greater than 1 and lower than 0) and the violation of the assumptions of normality and homoscedasticity of the errors distribution.

The solution to the boundary problem it is to find a more appropriate non-linear relationship, which would look like Figure 4.1, where the curve levels off and approaches the ceiling of 1 and the floor of 0. Conceptually, this S-shaped curve makes better sense than the straight line for representing the relationship when there is a dichotomous dependent variable.

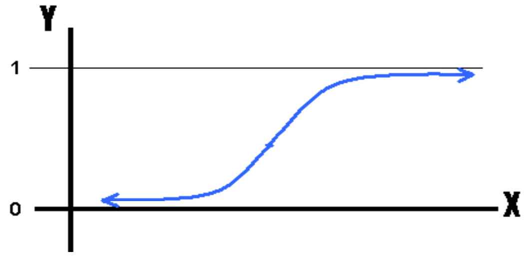

Figure 4.1 – The S-shaped Curve

4.3.5.1 Transforming Probabilities into Logits for Dichotomous Dependent variables

Linear regression faces a problem in dealing with a dependent variable with a ceiling and a floor that requires the same change in **X** to have a different effect in **Y** depending on how close the curve is to the maximum or minimum **Y** value. We need a transformation of the dependent variable to allow the decreasing effects of **X** on **Y** as the predicted **Y** value approaches the floor or ceiling.

Although many non-linear functions can represent the S-shaped curve, the "logistic" or "logit" transformation, because of its desirable properties and relative simplicity, has become popular [33]. To illustrate the logit transformation, assume that each case has a probability of having a characteristic or experiencing an event, defined as P_i. Since the dependent variable has values of only 0 and 1, this P_i must be estimated, but it helps to treat the outcome in term of probabilities for now. Given this probability, the logit transformation involves two steps. First, take the ratio of P_i to $1- P_i$, or the "odds" of experiencing the event. Second, take the natural logarithm of the odds. The logit thus is:

$$L_i = \ln [P_i / (1 - P_i)] \quad \text{……………………………………………} \quad (4.9)$$

For example, if P_i equals 0.2, its odds is (0.2 / 0.8) or 0.25, and its logit is ln (0.25) or -1.386. If P_i equals 0.7 , its odds is (0.7 / 0.3) or 2.33, and its logit is ln (2.33) or 0.847.

As seen, transforming probabilities into logit is easy in mathematical terms, but it requires some additional explanation to show its usefulness to describe the relationship between independent variables and a distribution of probabilities defined by a dichotomous dependent variable.

4.3.5.2 Meaning of the Odds

The logit begins by transforming probabilities into odds. Probabilities vary between 0 and 1, and express the likelihood of an event as a proportion of both occurrences and non-occurrences. Odds express the likelihood of an occurrence relative to the likelihood of a non-occurrence. Both probabilities and odds have a lower limit of zero, and both express the increasing likelihood of an event with increasingly large positive numbers, but they differ about the upper limit.

Unlike a probability, odds have no upper bound or ceiling. As a probability gets closer to 1, the numerator of the odds becomes larger relative to the denominator, and the odds become an increasingly large number. The odds thus increase greatly when the probabilities change only slightly near their upper boundary of 1. For example, probabilities of 0.99, 0.999, 0.9999, and so on result in odds of 99, 999, 9999, and so on. Tiny changes in probabilities results in huge changes in the odds, and that shows the odds increase toward infinity as the probabilities come closer and closer to 1.

Manipulating the formula for odds gives further insight into their relationship to probabilities. Beginning with the definition of odds ($\boldsymbol{O_i}$) as the ratio of the probability to one minus the probability, it is possible, with simple algebra, to express the probability in terms of odds:

$$\boldsymbol{P_i} \,/\, (\, 1 - \boldsymbol{P_i}) = \boldsymbol{O_i} \quad \text{implies that} \quad \boldsymbol{P_i} = \boldsymbol{O_i} \,/\, (\, 1 + \boldsymbol{O_i}\,) \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4.10)$$

The probability equals the odds divided by one plus the odds. Based on equation 4.10, the probability can never equal or exceed one; no matter how large the odds become in the numerator, they will always be smaller by one than the denominator. Of course, as the odds become large, the gap between the odds and the odds plus 1 will become relatively small and the probability will approach, but not reach, one. Conversely, the probability can be zero but it can never fall bellow zero. As long the odds equal or exceed zero, the probability must equal or exceed zero.

Usually, the odds are expressed as a single number, taken implicitly as a ratio to 1. Thus, odds of 10 imply an event will occur 10 times for each time it does not occur. The odds of 7 to 3 can be expressed equally well as a single number of 2.33 (to 1). Odds less than 1 mean the event is less likely to occur than it is to not occur. For example, if the probability is 0.3, the odds are 0.3 / 0.7 or 0.429. This means the event occurs 0.429 times per each time it does not occur. It could also be expressed as 42.9 occurrences per 100 non-occurrences.

In summary, creating odds represents the first step of the logit transformation. Also, reliance on odds rather than probabilities provides a useful interpretation of the likelihood of events. In addition, odds solve or eliminate the problem of upper boundary (ceiling) because they insure that the probability can never equal or exceed one.

4.3.5.3 Natural Logarithm of Odds

As mentioned previously, when using odds the probability can never fall below zero but it can equal zero. Then, it is necessary to solve the problems of the lower boundary equal to zero.

Taking the natural log of the odds eliminates the floor of zero. The natural log of:

- Odds above 0 and below 1 produces negatives numbers;
- Odds equal to 1 produces 0; and
- Odds above 1 produce positive numbers.

The natural logs of values equal to or below zero do not exist.

Logit transformation has some important properties or characteristics. Unlike a probability, the logit has no upper or lower boundary. The odds eliminate the upper boundary of probabilities, and the logged odds eliminates the lower boundary of probabilities as well. To see this, if P_i = 1, the logit is undefined because the odds of 1/0 do not exist. As the probability comes closer and closer to 1, the logit moves toward positive infinity. P_i = 0, the logit is undefined because the log of the odds of 0/1 or 0 do not exist. As the probability comes closer and closer to 0, the logit moves toward negative infinity. Thus, the logits vary from negative infinity to positive infinity. The problem of a ceiling and floor in the probabilities disappears.

Other properties are that the logit transformation is symmetric around the midpoint probability of 0.5. In addition, the same change in probabilities translates into different changes in the logits. These principles are shown in the following table:

Table 4.1 – Values of Probabilities, Odds, and Logits

P_i	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$1 - P_i$	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
Odds	0.111	0.25	0.429	0.667	1	1.5	2.33	4	9
Logit	-2.20	-1.39	-0.847	-0.405	0	0.405	0.847	1.39	2.20

A change in probabilities of 0.1 from 0.5 to 0.6, or 0.4 to 0.5, results in a change of 0.405 in the logit. But, the same probability change of 0.1 from 0.8 to 0.9, or 0.1 to 0.2, results in a change of 0.810 in the logit. The change in the logit for the same change in the probability is twice as large at this extreme as in the middle. Then the general principle is that small differences in probabilities result in increasingly larger differences in logits when the probabilities are near the bounds of 0 and 1.

It helps to view the logit transformation as linearizing the inherently nonlinear relationship between **X** and the probability of **Y**. Then, the linear relationship between the independent variables and the logged odds is:

$$\ln[P/(1-P)] = a + B_1 X_1 + B_2 X_2 + \ldots\ldots\ldots B_n X_n \qquad (4.11)$$

Where P is the probability of **Y**=1. Note that equation 4.11 implies non-linear relationship between the independent variables (X_1, X_2, … X_n) and probability. This equation is the basis of the logistic regression model.

4.4. The Logistic Regression Model

To express the probabilities rather than the logit as a function of the independent variables, each side of the equation 4.11 is taken as an exponent:

$$[P/(1-P)] = \exp a * \exp B_1 X_1 * \exp B_2 X_2 * \ldots\ldots\ldots \exp B_n X_n \qquad (4.12)$$

Thus, the odds change as a function of the coefficients treated as exponents. Solving equation 4.12 for P :

$$P = 1/[1 + \exp(-a - B_1 X_1 - B_2 X_2 - \ldots B_n X_n)] \ldots\ldots\ldots\ldots\ldots\ldots \qquad (4.13)$$

This equation shows the calculation of probabilities as a function of the independent variables and the coefficients. For purposes of calculations, it is the simplest form of logistic regression model.

4.4.1 Calculation of the Logistic Regression Coefficients

Equation 4.13 ensures that the probabilities estimated will not be greater than 1 or lower than 0, but because the linear form of the model (Equation 4.11) has infinitely large or small values of the dependent variable (logged odds), the method of ordinary least square (OLS) can not be used to estimate the coefficients in logistic regression. Instead, *maximum likelihood* techniques are used to calculate these parameters. The *maximum likelihood estimation* (MLE) maximizes the value of a function, the *log-likelihood* function, which indicates how likely it is to obtain the observed values of **Y**, given the values of the independent variables and the coefficients $\boldsymbol{a}$, $\mathbf{B_1}$, $\mathbf{B_2}$,… $\mathbf{B_n}$. Unlike the OLS method, which is able to calculate the coefficients directly, the solution or calculation of the coefficients for logistic regression models is found by beginning with a tentative solution, and iterating until a change in the likelihood function form one step to another is negligible.

The likelihood function (**L**) measures the probability of observing the particular set of dependent variable values ($\mathbf{P_1}$, $\mathbf{P_2}$,…$\mathbf{P_n}$) that occurs in the sample. This is written as the probability of the product of the dependent variable.

$\mathbf{L}$ = Prob (P_1 * P_2 * …P_n)

The higher the likelihood function, the higher the probability of observing the **P**s in the sample. The objective of the MLE technique is finding the coefficients ($\boldsymbol{a}$, $\mathbf{B_1}$, $\mathbf{B_2}$,… $\mathbf{B_n}$) that makes the log of the likelihood function (log **L**, usually named **LL**) as large as possible or -2 times the log of the likelihood function (**-2LL**) as small as possible.

The procedure is done by means of computer-implemented numerical algorithms designed to search for and identify the best set of parameters to maximize the log-likelihood function. In this research the “R” computer software was applied for calculation of the coefficients of the logistic models. This software is available from GNU General Public License, 1991, Free Software Foundation, Inc.

4.4.2 Interpretation of the Logistic Regression Coefficients

The interpretation of the logistic regression coefficients is not as straightforward as for the linear regression coefficients. The effects of the independent variables in a logistic regression have multiple interpretations. Effects exist for probabilities, odds, and logged odds and the interpretation of each effect have both advantages and disadvantages.

The effects of the independent variables on the logged odds are linear and additive (see Equation 4.11), but the units of the dependent variable, logged odds, have little intuitive meaning. The effects of the independent variables on the probability have intuitive meaning, but are non-linear and non-additive (see Equation 4.13). Despite the interpretable units, the effects on probabilities cannot be simply summarized in the form of a single coefficient.

The interpretation of the effects of the independent variables on the odds offers a compromise between the previous alternatives. The odds have more intuitive meaning than logged odds, and can express effects in single coefficients. The effects on odds are multiplicative rather than additive (see Equation 4.17) and the coefficients are exponential, but still have a straightforward interpretation. For the exponential coefficients, then, a coefficient of 1 leaves the odds unchanged, a coefficient greater than 1 increases the odds and a coefficient smaller than 1 decreases the odds. Moreover, the greater or lower the coefficient from 1, the greater the effect in changing the odds. In other words, the distance of an exponential coefficient from 1 indicates the size of the effects on odds. Then, the percentage increase or decrease of odds (%Δ) due to a one-unit change in the independent variable is:

$$\%\Delta = (\exp B - 1) * 100$$

4.4.2.1 The Baysian Information Criterion

The Baysian information criterion (BIC) developed by Raftery (1995) and cited by Pampel[33], is applied to each individual logistic regression coefficient in order to evaluate how the coefficient and its respective independent variable affect the logistic model. The BIC is defined as:

$$BIC = Z^2 - Ln(n) \quad \text{(4.14)}$$

Where Z is the ratio of the coefficient divided by its standard error, and "n" is the sample size. The value of BIC should exceed zero to reach significance. If the BOC value for a variable equals or falls bellow zero, the data provide little support for including the variable in the model. Raftery specifies grades of significance when the BIC is greater than zero. He defines BIC ranges from 0 to 2 as "weak", 2 to 6 as "positive", 6 to 10 as "strong", and greater than 10 as "very strong". This concept provides a helpful tool for logistic regression, where the interpretation of the coefficients is not easy.

4.4.3 Evaluation of Logistic Regression Models

There are several statistical methods for evaluating the performance of logistic models. For this research, the "Percent Correct Predictions" method was used to evaluate the models. This method assumes that if the estimated probability, **P**, is greater than or equal to 0.5 then the event is expected to occur and not occur otherwise. By assigning these probabilities, 0s and 1s, and comparing these to the actual 0s and 1s, the % correct Yes, % correct No, and overall % correct scores are calculated. Table 4.2 shows the results of the evaluation of a logistic regression model using the technique of percent correct prediction.

Table 4.2. Example of evaluation of a logistic model using the Percent Correct Predictions				
MODEL EVALUATION				
		PREDICTED		**% CORRECT**
		0	**1**	
OBSERVED	**0**	**210**	**41**	**83.67%**
	1	**19**	**1032**	**98.19%**
OVERALL				**95.39%**

4.5 Conclusions and Observations

The principles and derivation of logistic models, as well as the calculation and interpretation of their coefficients were explained in a simple form.

In addition, the evaluation of logistic models using the "Percent Correct Predictions" method was also illustrated.

The Baysian Information Criterion (BIC) and the percent correct prediction method will be used in the next chapters of this study for evaluating the grade of significance of the independent variables and the performance of logistic models respectively.

5 APPLICATION OF LOGISTIC REGRESSION TO DIAGNOSE BIT PERFORMANCE

5.1 Introduction

In response to the shortcomings of the baseline method, the statistical binary logistic regression models were used as an alternative way to diagnose bit performance. The models were developed taking into account as "independent variables" the drilling, conventional, and normalized parameters such as ROP, WOB, torque, specific energy, apparent formation strength, and the new parameters (F, G) proposed by Smith and Aghassi [1,37]. Initially, the balled bit condition was used as a "dependent dichotomous variable" with YES = 1 if the bit is balled at certain depth or time, and NO = 0. Then, the strong rock occurrence was taken as a second separate dependent variable with YES = 1 if the formation drilled is strong. The outcomes or probabilities of these two analyses were used as a tool to determine which of the two events, balled bit or strong rock, is more likely to be occurring.

In this regression, the outcome (probability that an event occurs) is measured on a binary scale between zero and one. Then, if the result is one or close to one, the event (e.g. bit balling) is true. On the other hand, if the result is zero or close to zero, the outcome indicates that the event is false. This technique allows the use of any number of independent variables (parameters) for developing the model, but for every independent variable it is necessary to calculate its respective coefficient. In order to calculate the coefficients, this method requires a training data set from an interval with known drilling parameters where the event or situation (e.g. bit balling or strong rock) in study has already happened. The greater the number of independent parameters used to develop the model, the greater the number of coefficients, and the more difficult it is to calculate the coefficients. Therefore, three or four independent parameters are a good practical limit for developing the logistic model for predicting strong rock or a balled bit. Once the coefficients are calculated from numerical methods and computer statistical software, the logistic model developed can be applied for evaluating the respective event or situation in study.

The logistic regression was applied to the data from wells from the Gulf of Mexico, Oklahoma, and South America. For each well data studied, several logistic models with three different independent variables were analyzed. The sets of three independent variables used were: the drilling parameters (WOB, ROP, Torque), the conventional parameters (Torque/WOB, Torque/ROP, ROP/WOB), the normalized parameters (ES, RF, FORS), and the new diagnostic parameters (F, G, dWOB/dt). Then, the best model for each situation was chosen using multiple criteria such as statistical evaluation (percent correct prediction for global model, and BIC factor for independent variables, see Equation 4.14) and conclusive response in known conditions.

The results from two wells will be described herein. First, the bit run # 12 from Matagorda Island Well # 6 is used to illustrate and evaluate the bit balling prediction. Second, the bit run # 15 from Matagorda Island Well # 1, where the baseline method diagnosed strong rock as bit balling under certain conditions (see chapter 3, section 3.2), is utilized to evaluate how useful the logistic regression is for overcoming this shortcoming.

5.2. Logistic Regression for Diagnosing Bit Balling

The down-hole data from Matagorda Island Well # 6, bit run # 12, over the interval 12,200'-12,900', was used to test the logistic regression as a tool to diagnose bit balling. The data set is described in chapter 3, section 3.1.1 and Figure 3.1. The interval 12,250'-12,380' was selected as a training data set for calculating the logistic regression model's coefficients. There are two shale sections, 12,250'-12,275' and 12,300-12,350', in this interval with strong, conclusive indications of a balled bit. The rest of the training interval, 12,275'-12,300', is interbedded sand and shale with a relatively clean drill bit. As a consequence, the "dependent dichotomous variable" for whether the bit is balled, was defined as follow for the intervals of the training data set:

- 12,250' – 12,275': YES = 1, balled bit
- 12,275' – 12,300': NO = 0, clean bit
- 12,300' – 12,350': YES = 1, balled bit

In order to developed a logistic model, it is also necessary to have the respective training data containing the independent variables. Three models were developed taking the following set of independent variables:

1) Conventional parameters:

- TDW = Torque / WOB
- TDR = Torque / ROP
- RDW = ROP / WOB

2) Normalized parameters:

- FORS: as defined in Equation 2.2
- ES: as defined in Equation 2.3
- RF: as defined in Equation 2.4

3) New derivative parameters (F and G) proposed by Aghassi and Smith [1,37] and dWOB/dt:

- F = f (Torque, WOB)
- G = g (ROP, WOB)
- D = dWOB / dt

D = dWOB/dt is not a diagnostic parameter. It is an operational input controlled by the driller in the field. Nevertheless, it could be potentially useful, especially during the "drilling off" periods, for correlation with other parameters for diagnosing bit balling, strong and weak rocks. For this reason, this parameter is included as a independent variable in logistic models.

In order to calculate the values of the models' coefficients, the "R" statistical software was applied using the training data sets. The results and the respective logistic model for each situation are shown in table 5.1.

Table 5.1- Logistic Models for Diagnosing a "Balled bit" using Matagorda Island Well # 6 Data with Different Parameters as Independent Variables

Independent variables	Coefficients		Logistic Model
Conventional parameters	Intercept	3.162	$P = \dfrac{1}{1+e^{-(3.162-22.498*TDW+16.791*TDR+0.337*RDW)}}$
	TDW	-22.498	
	TDR	16.791	
	RDW	0.337	
Normalized parameters	Intercept	-0.306	$P = \dfrac{1}{1+e^{-(-0.306+0.125*FORS-0.034*ES-0.751*RF)}}$
	FORS	0.125	
	ES	-0.034	
	RF	-0.751	
New parameters	Intercept	2.786	$P = \dfrac{1}{1+e^{-(2.786+5.6E-05*F-6.103*G+2.12*D)}}$
	F	2.56E-05	
	G	-6.103	
	D	2.12	

5.2.1. Evaluation of the Models

The evaluation of the coefficients based on the Baysian Information Criterion (BIC, see section 4.4.2.1), is illustrated in Table 5.2. From these results, it can be determined that most of the coefficients have a low value of "deviation error" compared to their value. As a result, high values of "Z" and "BIC" are obtained. As explained in Chapter 4, values of BIC equal or lower than zero means that the data provide little support and the respective independent variable should not be included in the model. For values greater than zero, the level of significance is classified as weak (0-2), positive (2-6), strong (6-10), and very strong (greater than 10). Then, statistically, the model with conventional parameters has the best resolution because all the three independent variables (TDW, TDR, RDW) have a very strong grade of significance. On the other hand, for the model with normalized parameters, the intercept is not significant and the force ratio (RF) has a weak impact in the model. Conversely, the two remaining normalized parameters (FORS and ES) have a strong impact in the model. In the third logistic model (new parameters), F, G, and D have a strong, positive and very strong grade of significance respectively.

Table 5.2 Evaluations of Coefficients of Logistic Models for Matagorda Island Well # 6. Sample size **n** = 1301.

Independent variables	**Coefficients**		**Standard error**	$Z = \frac{Coefficient}{S.error}$	**BIC = Z^2 - ln(n)**	**Grade of significance**
Conventional parameters	Intercept	3.162	0.4401	7.18	44.45	Very strong
	TDW	-22.498	2.0694	-10.87	111.02	Very strong
	TDR	16.791	1.9506	8.61	66.93	Very strong
	RDW	0.337	0.0563	5.99	28.66	Very strong
Normalized parameters	Intercept	-0.306	0.4833	-0.63	-6.77	Not significant
	FORS	0.125	0.0183	6.83	39.49	Very strong
	ES	-0.034	0.0082	-4.15	10.02	Very strong
	RF	-0.751	0.2744	-2.74	0.32	Weak
New parameters	Intercept	2.786	1.55E-01	17.96	315.48	Very strong
	F	2.56E-05	6.74E-06	3.80	7.26	Strong
	G	-6.103	1.901	-3.21	3.14	Positive
	D	2.12	2.01E-01	10.54	103.85	Very strong

A preliminary check of the effectiveness of the models for diagnosing bit balling is shown in Figure 5.1 where the training data is compared with each models prediction. In addition, Table 5.3 shows the results of the "Percent correct predictions" method, explained in section 4.4.3, for each model. As seen in Figure 5.1 and Table 5.3, all the models provide a reasonable diagnosis of the balling situation for the training interval with a percentage of correct prediction higher than 90%. Two of the models, conventional and normalized parameters, have almost the same tendency and can detect the intermediate shale located at 12,283'-12,287' where the ROP dropped from 76 ft/hr to 27 ft/hr. This situation is registered as apparent bit balling of very short duration with a probability of occurrence of about 90%. The model calculated with the new diagnostic parameters has the lowest percentage of correct predictions, 93.7%, which is also good. This model also detects the intermediate shale, however the diagnosis of this model seems to be delayed about 5 feet off depth. That happens because, as explained in section 3.1.4, the calculation of the new diagnostic parameters is made taking a linear regression with the previous 50 points (5 feet) of drilling parameters. For this reason, the response of this model is 5 feet late, which reduces the percentage of correct prediction by the model.

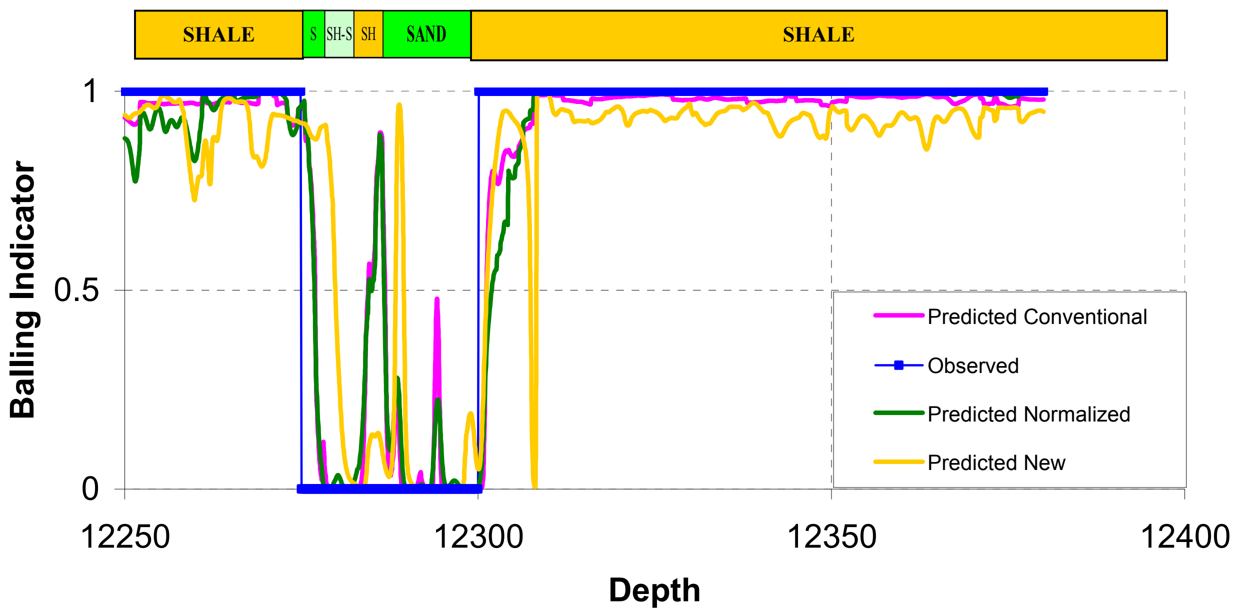

Figure 5.1. Plots of Logistic Models for Predicting Bit Balling in Matagorda Island Well # 6 .

Table 5.3. Evaluation of Logistic Regression Models for Predicting Bit Balling in Matagorda Island Well # 6										
EVALUATION		**Conventional Parameters**			**Normalized Parameters**			**New Parameters**		
		PREDICTED		**% CORRECT**	**PREDICTED**		**% CORRECT**	**PREDICTED**		**% CORRECT**
		0	**1**		**0**	**1**		**0**	**1**	
OBSERVED	**0**	**207**	**44**	**82.47%**	**210**	**41**	**83.67%**	**190**	**61**	**75.70%**
	1	**11**	**1040**	**98.95%**	**19**	**1032**	**98.19%**	**20**	**1030**	**98.10%**
OVERALL		**95.78%**			**95.39%**			**93.77%**		

In summary, the logistic model developed with conventional parameters is the best model because it has the highest percent of correct predictions and all of its independent variables have a very strong impact in the model. In addition, it responds very well to the known conditions, such as lithology change and a balled bit. However, for this example, the statistical evaluation and the response in known conditions of all three models are good because all the models have percent of correct predictions greater than 90%. In addition, all of their independent variables have some or a strong grade of significance. Then, all of the logistic models match well with the training conditions for diagnosing bit balling.

After getting the evidence that logistic regression models can potentially provide a good diagnosis for the balled bit condition, the procedure was tested for detecting strong rock conditions. The results and models obtained are described in the next section.

5.3 Logistic Regression for Distinguishing between Strong Rock and Bit Balled

As seen in section 3.2, Figure 3.6 and Table 3.12, the baseline method often diagnoses strong rock as a balled bit depending on the selection of the baseline zone. The logistic regression is potentially a tool to overcome this shortcoming. In order to evaluate the logistic regression as a means to distinguish between a balled bit and strong rock, the same data from Matagorda Island well # 1 (interval 13,200'-14,500') described in section 3.2.2 and Figure 3.5 was used to develop the logistic models to diagnose bit balling and strong rock independently. Then, as done in section 5.2 for Matagorda Island well # 6, one training data set is selected to generate a logistic model to diagnose bit balling. In addition, another training data set is selected to generate another logistic model to diagnose strong rock. After applying and interpreting these models independently, it maybe possible to determine which of the two situations, a balled bit or strong rock, is more likely to be occurring.

As seen previously in chapter 3, the Matagorda Island Well # 1 has several shale intervals with evidence of a balled bit, one of them at 13,460'–13,520'. This zone of low ROP is then followed by an interval of sand (13,522'-13,600') with high ROP and apparently a clean bit.

Then, the "dependent dichotomous variable" whether the bit is balled, was defined as follows for the intervals of the training data set:

- 13,470' –13,513': YES = 1, balled bit
- 13,530' – 13,600': NO = 0, clean bit

Conventional, normalized, and new diagnostic parameters were used as independent variables to develop logistic models for a balled bit using the statistical "R" software. After evaluating the models statistically, the best resolution was obtained with normalized parameters. The results for this model are shown in Table 5.4. As seen in this table, statistically, the model has a high responsiveness because the coefficients calculated have relatively small standard error and "very strong" BIC.

Table 5.4. Evaluations of the Bit balling Logistic Model's Coefficients for Matagorda Island Well # 1. Sample size **n** = 1601.

Independent variables	Coefficients		Standard error	$Z = \frac{Coefficient}{S.error}$	BIC = Z^2 - ln(n)	Grade of significance
Normalized parameters	Intercept	55.612	6.1	9.12	75.74	Very strong
	FORS	-0.8373	0.13668	-6.13	30.15	Very strong
	ES	0.4863	0.09119	5.33	21.06	Very strong
	RF	-66.4	7.7859	-8.53	65.35	Very strong
Logistic Model	$P = \frac{1}{1+e^{-(55.612-0.8373*FORS+0.4863*ES-66.40*RF)}}$					

This well has two intervals of strong rock, 13,310'-13,330' and 13,613'-13,645'. Now, the "dependent dichotomous variable", whether the formation is strong, is defined as follows for the intervals of the training data set:

- 13,635' – 13,641': YES = 0.9, strong rock.
- 13,265' – 13,270': NO = 0.1, weak rock.

There are several reasons for using YES=0.9 and NO=0.1 instead of YES=1.0 and NO=0. First of all, although the interval has two section of strong rock, these are not infinitely stronger as would be necessary to define YES=1. The same situation occurs for weak rock which is defined as NO=0.1. Second, when using the extremes values of YES=1 and NO=0 for calculations, the model may have a convergence problem, and the coefficients have such high deviation error that the BIC is negative resulting in a weak model. Taking into account that the logistic regression model allows establishing grades of occurrence or probability for the dichotomous dependent variable with values between zero and one, less extreme values can be used for the training data set in order to get a more realistic model.

The best model for "strong rock" was also obtained when using normalized parameters. The results from the "R" software and the respective model are shown in Table 5.5, where P_{SR} means probability of occurrence of strong rock.

Table 5.5. Evaluations of the Strong Rock Logistic Model's Coefficients for Matagorda Island Well #1. Sample size **n** = 112.

Independent variables	**Coefficients**		**Standard error**	$Z = \frac{Coefficient}{S.error}$	**BIC = Z^2 - ln(n)**	**Grade of significance**
Normalized parameters	Intercept	6.3397	1.436	4.41	14.77	Very strong
	FORS	-0.01547	0.00733	-2.11	-0.27	Not Significant
	ES	-0.00361	0.001614	-2.23	0.27	Weak
	RF	-3.38202	1.0977	-3.08	4.77	Positive
Logistic Model	$P_{SR} = \frac{1}{1+e^{-(55.612-0.8373*FORS+0.4863*ES-66.40*RF)}}$					

As seen in Table 5.5, the standard deviation is almost equal to the respective coefficients. As a consequence, low values of BIC are obtained for all the parameters except RF. The coefficient for ES has a weak grade of significance. The parameter FORS is not significant in the model, in contrast the variable RF has the highest level of significance for the “strong rock” logistic model.

The technique I propose to distinguish between strong rock and bit balling using logistic models is as follows. Once the best logistic models for diagnosing strong rock and bit balling are developed, these two models are applied independently. In the zones diagnosed as both bit balling and strong rock, one of them generally has more marked tendency to the value of probability = 1, so this is the more likely occurrence. Figure 5.2 illustrates the results of applying both models independently to the interval 13’300’-13,400’ of Matagorda Island Well # 1, where the baseline method proposed by Aghassi diagnosed the strong rock section as bit balled. This interval was not used in the training data for either model. In Figure 5.2, the probability that balling has occurred is plotted at the top in red, and the probability that strong rock is being drilling is shown in the second plot in blue. The plot of logs and drilling parameters are also shown. For the entire interval in Figure 5.2, the strong rock probability is higher than the balling probability. Consequently, this zone is diagnosed as stronger rock, and this result agrees with the geological interpretation because in this zone a section of strong siltstone was drilled. Then, using this statistical technique it is possible to confirm what situation, bit balling or strong rock, is happening.

5.4. Conclusions

Logistic regression models were developed and tested for diagnosing bit balling and strong rock. Values greater than 90% were obtained when the models were evaluated against training intervals using the criterion of percent of correct prediction.

The best logistic models for Matagorda Island Well # 1 used the Normalized parameters (Rf, ES, FORS) as independent variables for both the strong rock and balled bit. For Matagorda

Island Well # 6, the best model for bit balling prediction was obtained when using the conventional parameters (Torque/WOB, Torque/ROP, ROP/WOB).

Applying logistic models for diagnosing bit balling and strong rock individually was tested as a means to determine which of these two situations is more likely to be occurring Based on the analysis of one example of applying the combination of these regression models, they are potentially useful for distinguishing between strong rock and a balled bit.

I hypothesize that the logistic models can be used as a complement to Aghassi's method in order to improve the reliability of the diagnosis. A complete example of applying Aghassi's method and logistic models for diagnosing bit performance will be illustrated in the next chapter.

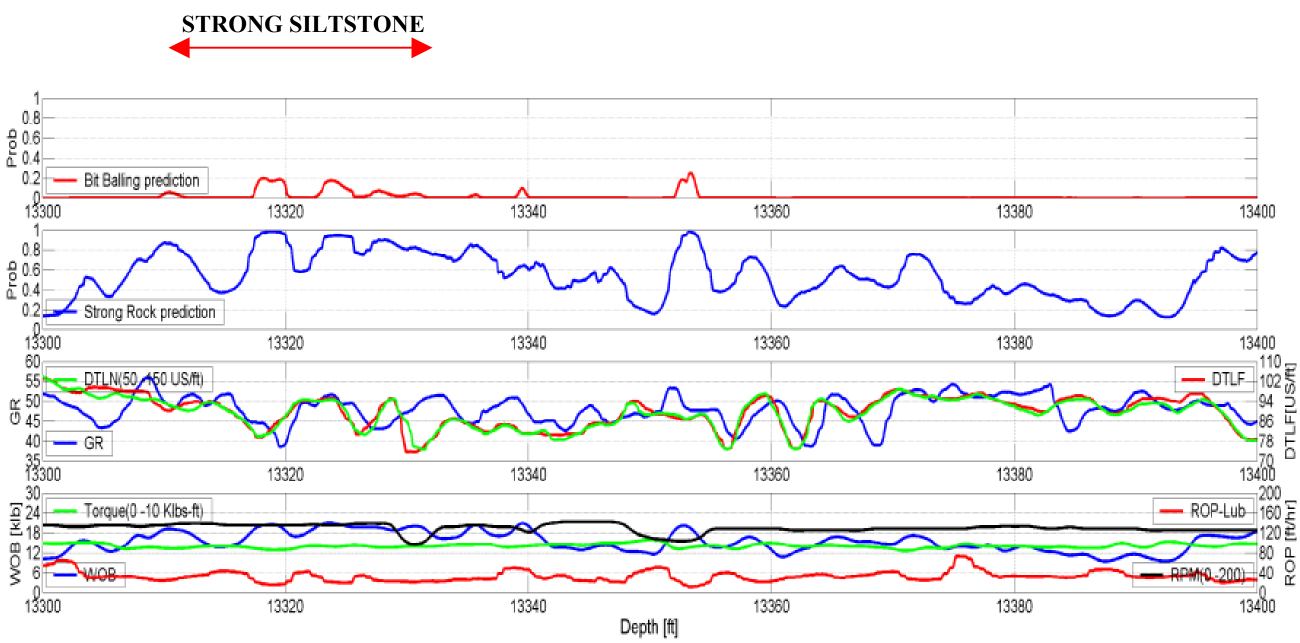

Figure 5.2 Logistic Regression Methods for Distinguishing a Balled Bit from Strong Rock. Matagorda Island Well # 1, Interval 13,300'-13,400'

6. LOGISTIC REGRESSION AS COMPLEMENT TO AGHASSI'S METHOD

6.1 Introduction

As seen in Chapter 3, the overall reliability of Aghassi's method when applied to drilling data from five wells was 51%. One of the most common causes of error was the wrong diagnosis of strong rock as a balled bit. In addition, the five diagnostic parameters of the method do not always agree conclusively about the diagnosis. For example, in some situations, the conventional parameters tend to predict bit balling, but the new diagnostic parameters tend to diagnose strong rock or a clean bit. For these reasons, I studied the logistic regression models, described in chapter 4, as a means of distinguishing between bit balling and strong rock situations. In Chapter 5, I developed independent logistic models to predict strong rock and bit balling in Matagorda Island Well # 1 (bit run # 15; 13,200'-14,500') where the strong rock sections were diagnosed as a balled bit with Aghassi's method. After applying the logistic models for predicting strong rock and bit balling independently, it was possible to determine which of these two situations were more likely to be occurring.

Consequently, logistic models can be used as either an alternative, or a complement, to Aghassi's method in order to overcome the shortcomings of that method in some specific situations. Then, the diagnosis can potentially be more accurate, and the reliability of the method potentially improved.

In this chapter, a complete example is evaluated by first applying Aghassi's method and then applying logistic regression models. The results were analyzed, evaluated and compared. The example involves four bits runs from a well in Oklahoma, which has very complete data including drilling parameters, electric logs, cores and sample analysis from an offset well with tests of ultimate stress, lithology interpretation, and daily drilling reports. Bit balling and strong rock situations were encountered in these bit runs. This conclusive data allows a good evaluation of the reliability of diagnosis using both Aghassi's method and logistic regression models for the known conditions.

After evaluating and comparing the results from both techniques, a procedure to apply Aghassi's method complemented with logistic regression models is described in this chapter. Finally, the effect of using logistic models as complement to Aghassi's method is evaluated.

6.2 Data Description – Oklahoma Well

This well was selected because high quality lithology data was available, and it was known to have experienced both bit balling and strong rock. The data analyzed is from four 12 ¼ inch PDC bit runs from 675' to 2,277'. An 8 bladed PDC bit, bit # 4, and then a 4 bladed PDC bit, bit # 5, were used for the four bit runs in the following intervals:

- Bit run # 4: 675'- 925'
- Bit run # 4R: 925' – 1,354'
- Bit run # 4RR: 1,354' – 1,975'
- Bit run # 5: 1,975'-2,277'

The drilling parameters were carefully measured in this well and recorded at about 20 second intervals. In addition, the well was logged with gamma ray, spontaneous potential, sonic, and caliper. Moreover, core samples were taken at different depths, and tests were performed in the lab to calculate rock properties such as Young's modulus, Poisson's ratios, yield stress and ultimate stress.

The geologic description provided shows that the lithology is mainly shale but also includes sand, siltstone, coal, and limestone. In addition, the plot of USC (uniaxial compressive strength) shows several sections of strong rock located about 1,750', 1,770', 1,820', and from 2,200' to the end of the well, where values of USC ranged from 25,000 psi to 50,000 psi. Moreover, the response of the drilling parameters describes two zones where it seems that the bit became balled. These zones are located at 1,500'-1,600' and 1,800'-1,975'. It is important to note that at the end of the bit run # 4RR at 1,975', the bit was pulled out because of the poor penetration rate while drilling a shale section. According to the drilling report, personal in the field observed that the bit was severely balled. In summary, the Oklahoma well presents an appropriate environment to apply and test the Aghassi's method and the logistic regression models.

6.3 Application of Aghassi's Method

Figure 6.1 describes graphically the response of the drilling parameters and the GR log for the four bit runs. In addition, the most relevant sand and strong intervals are also shown in this figure.

According to the behavior of the drilling parameters shown in figure 6.1, daily drilling reports, and the available information, the following operational interpretation can be made.

Bit run # 4 drilled from 675' to 925' with 12 Klbs WOB, 5 ft-klbs Torque, 120 RPM, and average ROP of 120 ft/hr. The section drilled was basically shale. After drilling 250' with good bit performance, the bit was pulled out at 925' to check its condition.

Bit # 4 was in good condition and was re-run (# 4R) from 925' to 1,350', drilling the long sand interval from 1,020' to 1,130'. The drilling parameters were adjusted according to bit performance and formation changes. For example, because ROP decreased from 120 to 80 ft/hr in the shale interval 925'-1020', the RPM were increased to 150 and WOB decreased to 9 Klbs.

When the bit entered the long sand interval at 1020', the ROP increased from 80 ft/hr to more than 120 ft/hr and Torque suddenly increased from 4 to 7 ft-Klb, then RPM were reduced from 150 to 120. At 1,350' the drill string was pulled out to check the bit because the ROP had dropped to 17 ft/hr. The bit was found to be in good condition with no obvious explanation for the decrease in performance.

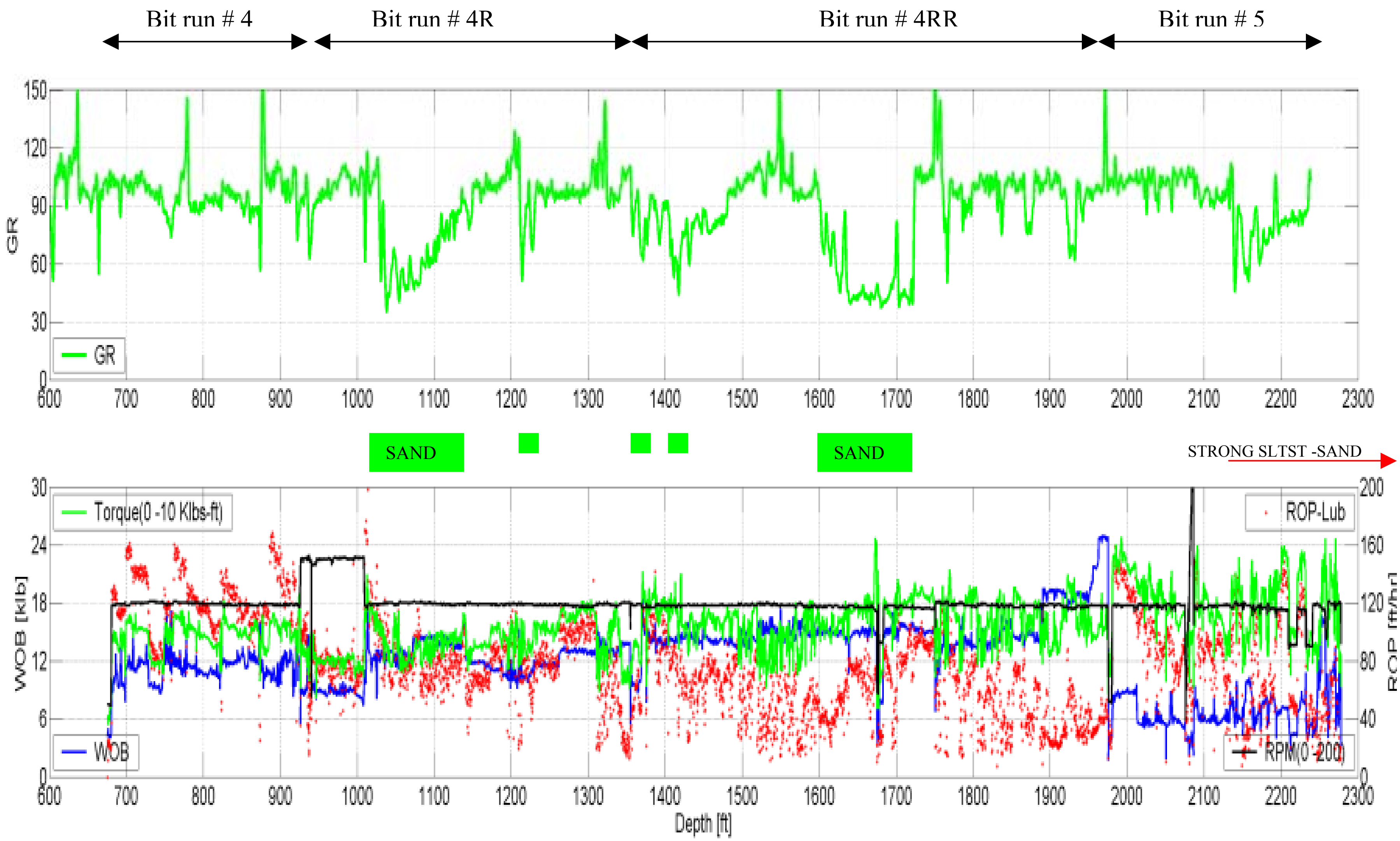

Figure 6.1. Drilling Parameters and GR log from Oklahoma Well.

Because the bit was in good condition, it was re-run again (# 4RR) at 1,350’ to 1,975’. A shale interval from 1,490’ to 1,600’ was drilled. At the end of this shale section, the ROP had a relatively low value of 40 ft/hr, and Torque decreased to 5 ft-klbs. The reduction in ROP and Torque can be indications that the bit was becoming balled in this interval. After that, a long interval of sand, 1,600’-1,710’, was drilled. After drilling the first 20’ of this sand, the ROP began to increase from 40 ft/hr to 80 ft/hr, possibly because the bit was cleaned by the sand. At 1,750’, the ROP began to vary from low values of 10 ft/hr to relatively high values of 80 ft/hr. This may be caused by the variation in formation types (siltstone, shale, limestone, clay, coal) observed in this section. About 1,900’ the ROP tended to stabilize at 30 ft/hr, and the WOB was increased to 20 klbs to improve bit performance. However, the ROP remained at relatively low values of 30-40 ft/hr. At 1,960’, the WOB was increased to a maximum value of 24 klbs, but the ROP did not respond, and Torque decreased from 7 to 5 ft-klbs. With this evidence of bit balling, the drill string was tripped out at 1,975’. The bit was observed balled when it was recovered, but it was not worn and was generally in good condition.

A new bit (# 5) was run from 1,975’ to 2,277’, drilling the strongest interval of this well located from 2,150’ to the end of the bit run. It is important to note that the large changes in RPM during this bit run, see Figure 6.1, are caused by several drilling tests that were performed in this section of the well.

6.3.1 Selection of the Baselines

One of the conclusions from Chapter 3 was, that besides Aghassi’s recommendation about locating the baseline in an interval of relatively high ROP over a long shale section, it is also important to locate the baseline in an interval of relatively high or average WOB. Therefore, both criteria were applied in this well. The first baseline, Baseline-1, was located at the beginning of the bit run # 4, in the shale interval 700’-730’, where an average WOB of 12 klbs was used. Because there were no drastic changes in drilling performance in about 1000’, this baseline will be used to evaluate the diagnostic parameters until 1,710’. Then, Baseline-1 will be the reference to evaluate the performance of bit runs # 4, # 4R, and part of # 4RR.

Because of changes in drilling response, lower ROP and greater Torque and WOB due to variation of rock properties in the formations drilled, the baseline was updated. The second baseline (Baseline-2) was located in the shale interval 1,710’-1,750’ with a relatively high WOB of 15 klbs and ROP of 90 ft/hr. This baseline will be the reference to evaluate the final section of bit run # 4RR and all of bit run # 5.

6.3.2 Calculation and Evaluation of the Diagnostic Parameters

The drilling data from the Oklahoma Well was available versus depth and versus time with time resolution in a range from 7 seconds to 20 seconds. Consequently, the diagnostic method was applied to the data versus time. For evaluating the method, the results were converted to depth in order to compare results with the lithology interpretation. The derivative parameters (F, G, and dWOB/dt) were calculated using linear regression with the 50 previous points.

Figure 6.2 shows plots of the drilling parameters (WOB, Torque, RPM, and corrected ROP), the conventional diagnostic parameters (Torque/WOB, Torque/ROP, ROP/WOB), and the new diagnostic parameters (F and G) versus depth for the four bit runs. Also, the Baseline-1 and Baseline-2 zones are shown in this figure.

The interval from 675’ to 2,277’ was evaluated as follows. After setting the baselines, the conventional and new diagnostic parameters are calculated and plotted, see Figure 6.2. Then the entire section (675’ – 2,277’) was divided into intervals by selecting the zones where there was a major change in the diagnostic parameters or known lithology change. Consequently, 21 intervals were evaluated in this well.

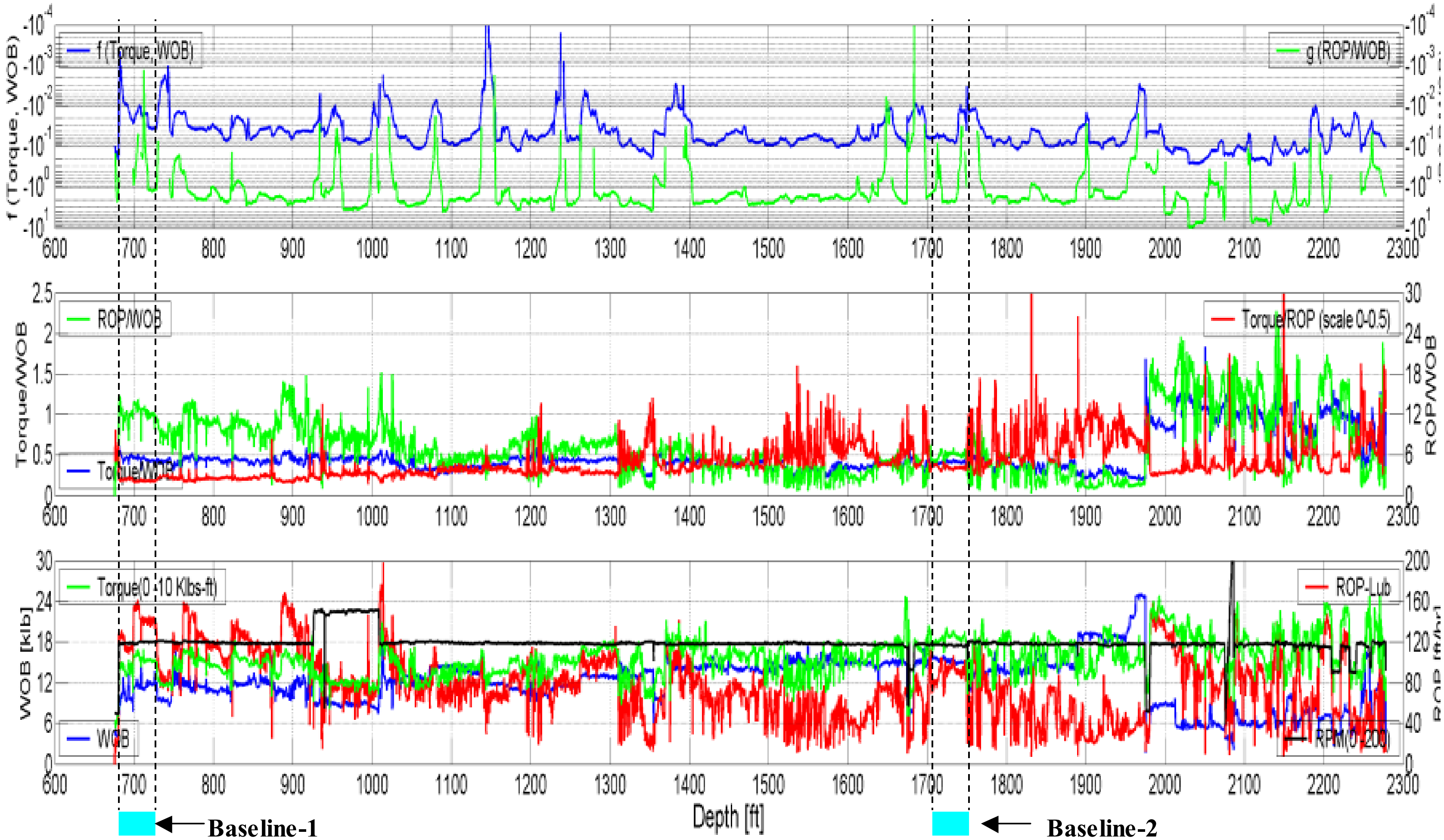

Figure 6.2 Conventional and New Diagnostic Parameters from Aghassi's Method Applied on Oklahoma Well

Appendix I is a table summarizing the interpretation of the diagnostic parameters for each of the 21 intervals studied. The lithology description from the offset well and the response of each diagnostic parameter are shown. As seen in Appendix I, 7 of the 21 intervals analyzed were correctly diagnosed and 3 were wrong. The 11 remaining intervals were partially diagnosed because they were correctly diagnosed by only one set of diagnostic parameters, conventional or new. In some intervals, the conventional parameters tended to predict bit balling, but the new diagnostic parameters tended to diagnose a clean bit or strong rock and vice versa. Therefore, the diagnosis for these intervals was not conclusive. Consequently, the efficiency of Aghassi's method for this well was 7 / 21 or 33%.

Specifically three intervals of shale were diagnosed as strong rock, and eleven were not clearly diagnosed because the method's parameters did not agree about defining these intervals as a balled bit, weak or strong rock.

In an attempt to improve the reliability of a method for diagnosing bit performance, logistic regression models were developed, applied, and evaluated as a complement to Aghassi's method in order to improve its diagnosis efficiency.

6.4. Logistic Regression Models for Oklahoma Well

Logistic regression models for diagnosis of bit performance were introduced in Chapter 5. Finding the best logistic model for predicting bit balling or strong rock is often difficult because of continuous variation in drilling parameters due to lithology changes. Therefore a detailed analysis must be made to select the best training data set and the independent variables with an acceptable grade of statistical significance for use in the model. Once the model is developed, it can be evaluated under known lithology conditions in order to determine how accurate the model is for predicting bit balling or strong rock.

The training data set and the diagnostic parameters used as independent variables in the model have to meet not only the statistical requirements (coefficients with low standard error and significance in the model), but also they have to predict correctly the known conditions (a balled bit, stronger rock, or neither) of the well in study. The procedure for developing a logistic model is as follow.

1) Considering the available drilling data, lithology interpretation, core evaluation, drilling reports, etc, select the training data sets for the situations to be diagnosed, bit balling or strong rock. For example, if the model is for predicting a balled bit, it is necessary to select at least two intervals, one where an extreme case of this situation exists and another where the opposite situation, a very clean bit, was evident. For each interval, it is necessary to assign the grade of severity of the event between 0 and 1, where 1 means severe bit balling and 0 means a clean bit. Extra intervals with intermediate grade of severity can be used as training data, and they have to be graded according to the severity observed. The same procedure is done if the model is for predicting strong rock, where at least one interval of stronger rock and one with weak rock have to be chosen and graded as a training data set. The grade of severity assigned is subjective and is used only as input data for the statistical software that determines the coefficients of the model.
2) Once the most significant training data is selected, the potentially relevant diagnostic parameters must be tested within groups. The grades of statistical significance (BIC, Equation 4.14) of their coefficients in the logistic model are calculated. Then, the models with independent variables that have the highest grades of significance are identified, and used to develop at least one model based on parameters with high significance to be tested against known data.

3) The best models selected from statistical analysis, are applied on the known lithology conditions in the training intervals and evaluated by the percentage of correct prediction technique. At the end, the model that combines both the best statistical significance and best response under known conditions is selected as the final model for predicting bit balling or strong rock.

6.4.1. Logistic Regression Model for Diagnosing Bit Balling

The first step in developing a logistic model is to define the training data or intervals that best represent the balling situation. According to the lithologic description, drilling performance, and daily reports, the bit run # 4RR was balled when the bit was tripped out at 1975'. This situation was corresponds with the interpretation of Aghassi's method in the interval 1950'-1975' which was diagnosed as balled bit in a shale section, see Appendix I. In addition, after changing the bit at that depth, the drilling performance of the new bit, #5, was considerably better than bit #4RR, because now the bit was clean. As a consequence, the interval of training data and the "dependent dichotomous variable" whether, the bit is balled were initially defined as follows:

1951.61' - 1974.93' = 1.0, balled bit

1980.63' - 2012.06' = 0, clean bit

Once the intervals of training data are defined, it is necessary to choose the parameters to be used as independent variables in the model. Initially, the basic drilling parameters (TORQUE, ROP, and WOB) were used as independent variables. But it was concluded, statistically, that they do not have significance in the models. Consequently, in order to develop the best logistic model for diagnosing bit balling, several alternatives were tested with conventional, normalized and new diagnostic parameters, changing or adjusting other conditions such as the training data intervals and the assigned grade of balling severity.

Table 6.1 shows the summary of the statistical evaluation of the coefficients for the 13 selected alternatives or models proposed for bit balling diagnosis. In this table, the following information is given for every model.

- The independent variables
- The training data intervals
- The size of the sample, **n**
- The value of the coefficients obtained from the "R" software for each independent variable
- The standard error for each coefficient estimated
- The value of Z = coefficient/standard error
- The value of BIC = Z – Ln (n), and
- The grade of significance according to Raftery's criterion as described by Pampel[33].

The models listed in Table 6.1 are in chronological order as they were formulated and analyzed. The most representative characteristic, or change, taken into account for developing every model is marked with a red circle line in Table 6.1.

As mentioned previously, the initial model, Model 1, was developed with the drilling parameters (ROP, TORQUE, WOB) as independent variables, but they do not have significance for the model because their BIC values are negative.

In Model 2, the grade of severity of balled (YES) and clean bit (NOT) were changed to 0.95 and 0.05 respectively, keeping the drilling parameters as independent variables. As seen in Table 6.1, after this modification the WOB has a strong grade of significance, but TORQUE and ROP still have negative values of BIC. Therefore, they are not significant for the model. In addition, WOB cannot be a sole diagnostic parameter in a model because it is the operational input parameter controlled by the driller.

Table 6.1 - Statistical Evaluation of Logistic Regression Models for Diagnosing Bit Balling in Oklahoma Well								
Model No	**Independent variables**	**Balling Training data**	**Coefficients**		**Standard error**	**Z**	**BIC**	**Grade of significance**
1	**Drilling parameters n (size sample) =**	1951.69-1974.93=1.0 1980.63-2012.06=0 167	Intercept	-8.0191	243.19	-0.03		
			ROP	-0.151	2.16	-0.07	-5.11	o
			WOB	0.9056	6.44	0.14	-5.10	o
			Torque	1.0915	64.43	0.02	-5.12	o
2	**Drilling parameters n (size sample) =**	1951.69-1974.93=0.95 1980.63-2012.06=0.05 167	Intercept	-8.5942	6.39	-1.34		
			ROP	0.0098	0.041	0.24	-5.06	o
			WOB	0.4915	0.132	3.72	8.75	Strong
			Torque	-0.0619	1.436	-0.04	-5.12	o
3	**Drilling parameters & BT n (size sample) =**	1648.92-1664.58=0.20 1951.69-1974.93=0.95 1980.63-2012.06=0.05 218	Intercept	-8.799	6.35E+00	-1.39		
			ROP	1.06E-02	4.00E-02	0.27	-5.31	o
			WOB	-0.0534	1.42	-0.04	-5.38	o
			Torque	0.4659	0.14	2.64	1.59	Weak
			BT	0.1389	3.60E-01	0.39	-5.24	o
4	**Conventional parameters n (size sample) =**	1648.92-1664.58=0.20 1951.69-1974.93=0.95 1980.63-2012.06=0.05 218	Intercept	-2.0032	2.84E+00	-0.71		
			TDW	-2.22E+01	8.91E+00	-2.49	0.83	Weak
			TDR	61.0139	16.32	3.74	8.59	Strong
			RDW	0.9242	3.60E-01	2.57	1.21	Weak
5	**Conventional parameters & BT n (size sample) =**	1648.92-1664.58=0.20 1951.69-1974.93=0.95 1980.63-2012.06=0.05 218	Intercept	-1.7661	3.20E+00	-0.55		
			TDW	-2.29E+01	9.95E+00	-2.30	-0.08	o
			TDR	62.5386	18.94	3.30	5.52	Positive
			RDW	0.9504	0.4	2.64	1.59	Weak
			BT	-0.0635	3.90E-01	-0.16	-5.36	o
6	**Normalized parameters n (size sample) =**	1648.92-1664.58=0.20 1951.69-1974.93=0.95 1980.63-2012.06=0.05 218	Intercept	-1.2852	2.66	-0.48		
			FORS	0.31	0.11	2.82	2.56	Positive
			ES	-0.09	0.1	-0.90	-4.57	o
			RF	-0.222	0.45	-0.49	-5.14	o
7	**(FORS)-1,ES.RF parameters n (size sample) =**	1648.92-1664.58=0.20 1951.69-1974.93=0.95 1980.63-2012.06=0.05 218	Intercept	-1.9082	2.8116	-0.68		
			1/FORS	45.8462	18.2316	2.51	0.94	Weak
			ES	0.1611	0.0423	3.81	9.12	Strong
			RF	-5.8174	2.3411	-2.48	0.79	Weak
8	**New parameters n (size sample) =**	1648.92-1664.58=0.20 1951.69-1974.93=0.95 1980.63-2012.06=0.05 218	Intercept	1.55	3.03E-01	5.12		
			F	39.63	8.1	4.89	18.55	Very Strong
			G	0.277	1.09E-01	2.54	1.07	Weak
			D	-1.90E-05	1.21E-05	-1.58	-2.90	o
9	**Conventional parameters & F n (size sample) =**	1648.92-1664.58=0.20 1951.69-1974.93=0.95 1980.63-2012.06=0.05 218	Intercept	-1.7447	3.42E+00	-0.51		
			TDW	-2.33E+01	1.20E+01	-1.95	-1.60	o
			TDR	60.1731	17.4251	3.45	6.54	Strong
			RDW	0.9622	0.4569	2.64	1.59	Weak
			F	-1.7787	1.28E+01	-0.14	-5.37	o
10	**F & G parameters n (size sample) =**	1648.92-1664.58=0.20 1951.69-1974.93=0.95 1980.63-2012.06=0.05 218	Intercept	1.3796	2.86E-01	4.83		
			F	43.251	8.3133	5.20	21.68	Very Strong
			G	0.1793	8.83E-02	2.03	-1.26	o

Table 6.1 (Continued)								
Model No	**Independent variables**	**Balling Training data**	**Coefficients**		**Standard error**	**Z**	**BIC**	**Grade of significance**
11	**Mix best parameters n (size sample) =**	1648.92-1664.58=0.20	Intercept	-3.4696	1.01E+00	-3.44		
		1951.69-1974.93=0.95	F	3.2709	9.0291	0.36	-5.25	o
		1980.63-2012.06=0.05	FORS	0.221	4.25E-02	5.20	21.66	Very Strong
		218	RDW	3.70E-03	7.45E-02	0.05	-5.38	o
12	**Mix best parameters n (size sample) =**	1648.92-1664.58=0.20	Intercept	-2.2938	1.86E+00	-1.23		
		1951.69-1974.93=0.95	FORS	0.305	0.1151	2.65	1.64	Weak
		1980.63-2012.06=0.05	TDR	-26.6724	3.69E+01	-0.72	-4.86	o
		218						
13	**Only FORS n (size sample) =**	1648.92-1664.58=0.20	Intercept	-3.62338	4.25E-01	-8.53		
		1951.69-1974.93=0.95	FORS	0.22566	0.02703	8.35	64.31	Very Strong
		1980.63-2012.06=0.05						
		218						

Two modifications were made for Model 3. First, the training interval 1648.92'–1664.58' with 0.2 grade of balled severity is included. It allows developing a model from more definitive training data, and it forces the model to have low values of balling prediction under certain conditions. In addition, the BT parameter is included as independent variable. This BT parameter is the B torque parameter proposed and later rejected by Aghassi[1] as a means to distinguish if a ROP reduction is due to a balled bit or strong rock. Including this modification in Model 3 supplied a worse result than Model 2. Only Torque has a "weak" grade of significance in the model.

In Model 4, only the conventional diagnostic parameters (TDW, TDR, RDW) are considered as independent variables with the same training intervals. As seen in Table 6.1, the conventional parameters created a potentially relevant model because all of the parameters have some grade of significance. TDW and RDW have a weak and TDR has a strong grade of significance.

In Model 5, the parameter BT is included to be tested with conventional parameters. Again, this parameter was not graded as significant in the model. In addition, it caused a lower grade of significance of the conventional parameters, negatively affecting the relevance of the model.

In Model 6, the normalized diagnostic parameters (FORS, ES, RF) are used as independent variables. The result was satisfactory only for FORS, which had a positive grade of significance in the model. Because normalized parameters are almost equivalent to conventional parameters, see section 2.5, it was expected that these results would be close to Model 4 where all of the parameters had significance, but it was not. That could happen because the conventional diagnostic parameter TDW is equivalent to 1/FORS, so they are not directly proportional. For this reason in Model 7, the normalized parameters were tested again but using 1/FORS instead of FORS as one of the independent variables. As seen in Table 6.1, the grades of significance of Model 7 variables are similar to Model 4 variables. TDW and 1/FORS as well as RDW and RF have a weak grade of significance, and TDR and ES have a strong grade of significance. Then, these models are essentially equivalent as expected.

In Model 8, the new diagnostic parameters (F and G) are used with "D" to develop the model. "D" is the term dWOB/dt already defined in chapter 5. Two parameters, F and G, have a very strong and weak grade of significance respectively in the model.

The statistical evaluation of the coefficients indicates that Model 4 (conventional parameters), Model 7 (Normalized parameters with 1/FORS) and Model 8 (new parameters) provided the best models for diagnosing bit balling. Then, several additional logistic models were developed and tested using the parameters with highest grade of significance (F, TDR, FORS, and RDW) as independent variables.

From the results obtained with Models 9, 10, 11, and 12 where four combinations of these parameters were tested, it was determined that the parameter FORS has such a strong significance in the model that it made the other parameters irrelevant for diagnosing a balled bit. For that reason in Model 13, only the parameter FORS was used as independent variable. The statistical evaluation of this model is very good, because the only parameter of the model has a "very strong" grade of significance.

In summary, from the statistical evaluation of the coefficients; Model 4, Model 7, Model 8 and Model 13 were the best for diagnosing bit balling. These four models were applied on the training data intervals in order to make a preliminary evaluation of the model using the "percentage of correct prediction" criterion. The results of that evaluation are shown in Table 6.2. Model 4, Model 7 and Model 13 have 100% correct prediction in the training intervals. Model 8 has an overall percentage of correct prediction of 78.5 %, which is lower but still high.

Table 6.2. Evaluation of Logistic Regression Models for Predicting Bit Balling in Oklahoma Well

EVALUATION		No 4- Conventional Par.			No7- ES, RF, 1/FORS			No 8-New Parameters			No 13- Only FORS		
		PREDICTED		% CORRECT	PREDICTED		% CORRECT	PREDICTED		% CORRECT	PREDICTED		% CORRECT
		0	1		0	1		0	1		0	1	
OBSERVED	0	137	0	100.00%	137	0	100.00%	105	32	76.64%	137	0	100.00%
	1	0	82	100.00%	0	82	100.00%	15	67	81.71%	0	82	100.00%
OVERALL		100.00%			100.00%			78.54%			100.00%		

In order to find the best model for balling diagnosis, the four models were applied to and evaluated in the entire well (675'-2277'). The response of each model is illustrated in Figure 6.3 where the drilling parameters and the GR plots are shown for visual evaluation of the models. As seen in this figure, Model 4 (conventional parameters), Model 7 (1/FORS, ES, RF), and Model 8 (new parameters), blue, black, and green curves in the top plot, appear to over-predict balling during the initial bit run (675'-925') which is not true because that interval was drilled with a ROP average of 120 ft/hr.

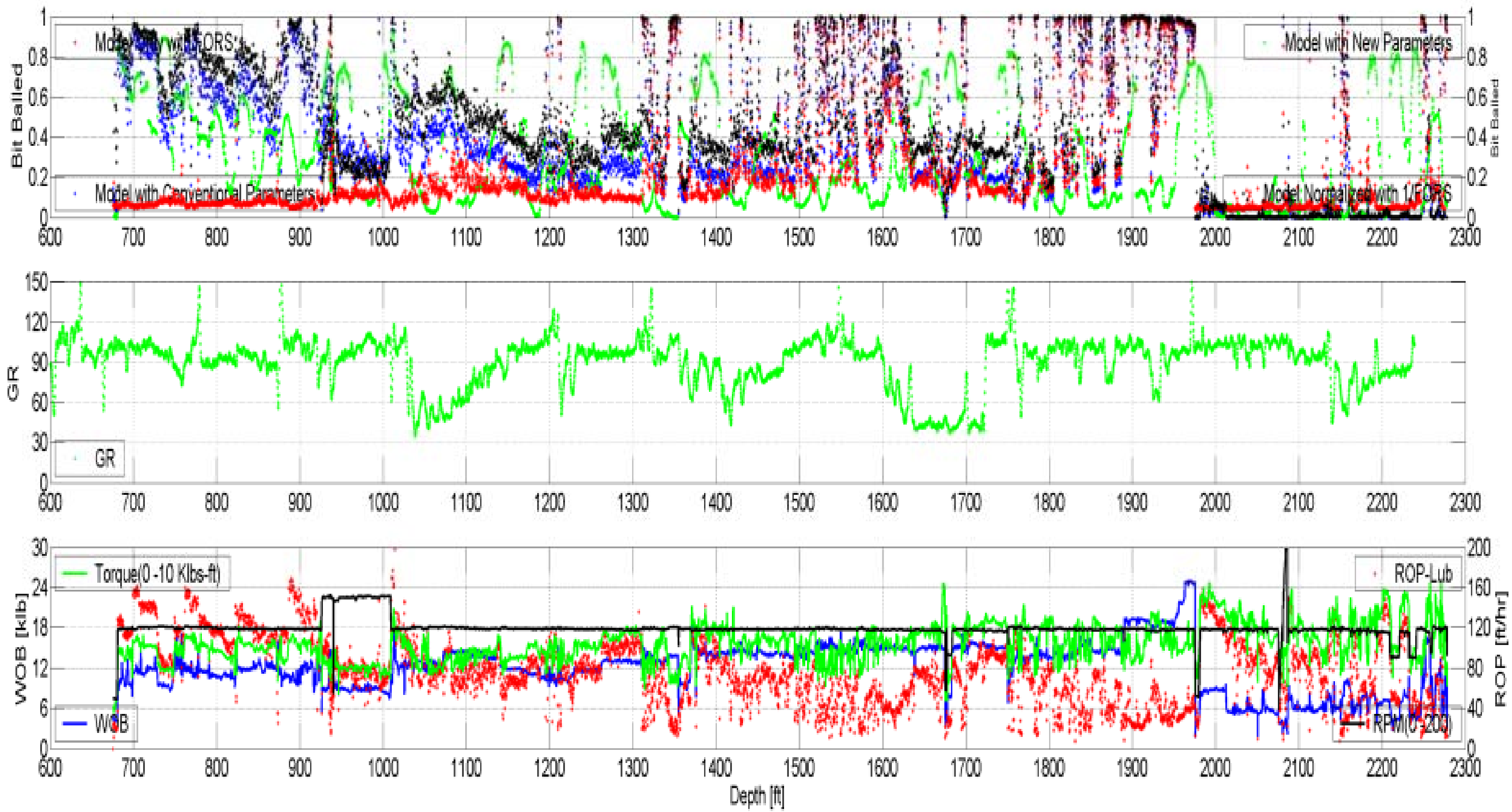

Figure 6.3 Logistic Regression Models for Diagnosing Bit Balling in Oklahoma Well

On the other hand, Model 13 (red curve) began to predict a balled bit at the end of bit run # 4R (1350'). This is probably correct, given the slow ROP in shale although the bit was clean when it was inspected after being pulled. In addition, it diagnoses balling in section of shale with relatively low ROP, which is reasonable.

In conclusion, Model 13 with only the normalized parameter FORS as an independent variable was selected as the best model for diagnosing bit balling in the Oklahoma Well. It is important to note that in Chapter 5, for Matagorda Island Well # 1, the best logistic model for bit balling prediction was obtained including the three normalized parameters (FORS, ES, RF) as independent variables, and all of them had a "very strong" significance. Then, these two models have a common independent variable, FORS, with very strong significance in the model.

6.4.2. Logistic Regression Model for Diagnosing Strong Rock

The same general procedure followed for the bit balling logistic models was used in order to find the best logistic regression model for diagnosing strong rock. According to the lithological description, drilling performance, and rock properties determined from cores and logging, strong intervals were drilled primarily during bit run # 5, 1,975'-2,277'. On the other hand, the weakest intervals were drilled during bit run # 4R (795'-1,350').

As a consequence, the intervals of training data and the "dependent dichotomous variable", whether the formation drilled was strong, were initially defined as follow:

2,231.91'-2,277.6' = 0.9, strong rock

940,05'-1000,19' = 0.1, weak rock

Initially, the basic drilling parameters (TORQUE, ROP, and WOB) were used as independent variables. Then, several alternatives were tested utilizing other parameters and conditions in order to develop the best logistic model for diagnosing strong rock. Table 6.3 shows the summary of the statistical evaluation of the coefficients for the 16 selected alternatives or models proposed for strong rock diagnosis. This table contains the same type of information as Table 6.2, and the models described in it are also in chronological order as they were formulated and analyzed.

The first 5 models were developed using the drilling, conventional, normalized, and new diagnostic parameters as independent variables. After the statistical evaluation of the coefficients, Model 3 (normalized parameters) and Model 5 (new parameters) are the best for strong rock diagnosis. In addition, the parameters with highest significance are ES, RF, F, and G. Consequently, several logistic models were developed and tested using combinations of these parameters with highest grade of significance as independent variables.

The results obtained in evaluating Models 6 to 12 demonstrated that the parameter ES and RF have the strongest significance in the model. For that reason, only the parameters ES and RF were used as independent variables in Model 13. The statistical evaluation of this model is good, because both of the parameters of the model have a "very strong" grade of significance.

A new parameter proposed by Smith [39] was tested as an independent variable for strong rock models. This derivative parameter, denoted as H in this study, was defined by Smith as d(Torque/ROP)/dWOB. It was tested with F, G, ES, and RF as independent variable. As seen in Table 6.3, the parameter H had a very strong significance in Model 14 when it was tested with F and G. In addition, H also had strong significance in Model 15 when it was tested alone. However, H had only a weak significance when it was tested in Model 16 with the parameters with strongest significance for strong rock (ES and RF).

In summary, from the statistical evaluation of the coefficients, Model 3, Model 5, Model 13, and Model 16 were the best for predicting strong rock. These 4 models were applied on the strong and weak rock training data intervals in order to make a preliminary evaluation of the model using the "percentage of correct prediction" criterion.

Table 6.3 - Statistical Evaluation of Logistic Regression Models for Diagnosing Strong Rock in Oklahoma Well

Model No	Independent variables	Strong Rock Training data	Coefficients		Standard error	Z	BIC	Grade of significance
1			Intercept	-16.0028	11.0517	-1.45		
	Drilling	2231.91-2277.6=0.9	ROP	-1.3451	0.6375	-2.11	-1.32	o
	parameters	940.05-1000.19=0.1	WOB	-3.1991	1.5133	-2.11	-1.31	o
	n (size sample)=	322	Torque	22.4541	10.3977	2.16	-1.11	o
2			Intercept	-3.1438	1.46E+00	-2.16		
	Conventional	2231.91-2277.6=0.90	TDW	6.79E+00	1.54E+00	4.40	13.57	Very Strong
	parameters	940.05-1000.19=0.10	TDR	15.3374	8.763	1.75	-2.71	o
	n (size sample)=	322	RDW	-0.3377	0.175	-1.93	-2.05	o
3			Intercept	-7.1378	0.9687	-7.37		
	Normalized	2231.91-2277.6=0.90	FORS	-0.07564	0.1491	-0.51	-5.52	o
	parameters	940.05-1000.19=0.10	ES	0.1195	0.0458	2.61	1.03	Weak
	n (size sample)=	322	RF	1.5451	0.477	3.24	4.72	Positive
4			Intercept	-5.498	1.57E+00	-3.51		
	Normalized	2231.91-2277.6=0.90	FORS	-8.98E-02	1.46E-01	-0.62	-5.39	o
	parameters & BT	940.05-1000.19=0.10	ES	0.12156	0.04455	2.73	1.67	Weak
	n (size sample)=	322	RF	1.2367	0.51428	2.64	1.20	Weak
			BT	-0.3296	2.60E-01	-1.27	-4.16	o
5			Intercept	-3.61E-01	2.66E-01	-1.36		
	New	2231.91-2277.6=0.90	F	-1.32E+01	5.118	-2.58	0.87	Weak
	parameters	940.05-1000.19=0.10	G	4.17E-01	7.73E-02	5.39	23.28	Very strong
	n (size sample)=	322	D	4.34E-06	3.45E-06	1.26	-4.20	o
6			Intercept	-6.68664	0.92077	-7.26		
	(FORS)-1, ES,RF	2231.91-2277.6=0.90	1/FORS	-8.68701	7.35567	-1.18	-4.38	o
	parameters	940.05-1000.19=0.10	ES	0.07454	0.02215	3.37	5.55	Positive
	n (size sample)=	322	RF	2.31016	0.53958	4.28	12.56	Very strong
7			Intercept	-4.48E-01	2.57E-01	-1.74		
	New	2231.91-2277.6=0.90	F	-1.42E+01	5.05751	-2.81	2.15	Positive
	parameters	940.05-1000.19=0.10	G	4.16E-01	7.74E-02	5.38	23.12	Very strong
	n (size sample)=	322						
8			Intercept	-6.89625	7.66E-01	-9.00		
	ES,RF, F, G	2231.91-2277.6=0.90	G	7.44E-02	8.43E-02	0.88	-5.00	o
	parameters	940.05-1000.19=0.10	F	6.40644	7.92665	0.81	-5.12	o
	n (size sample)=	322	ES	0.0967	0.01376	7.03	43.61	Very strong
			RF	1.67633	2.39E-01	7.03	43.62	Very strong
9			Intercept	-5.9854	1.38E+00	-4.35		
	Mix best	2231.91-2277.6=0.90	ES	0.09565	0.01338	7.15	45.33	Very strong
	parameters & BT	940.05-1000.19=0.10	RF	1.4977	3.12E-01	4.81	17.33	Very strong
	n (size sample)=	322	BT	-3.17E-01	2.59E-01	-1.22	-4.28	o
10			Intercept	-7.0372	7.55E-01	-9.32		
	Mix best	2231.91-2277.6=0.90	F	8.931	7.34	1.22	-4.29	o
	parameters	940.05-1000.19=0.10	ES	0.1005	1.33E-02	7.56	51.32	Very strong
	n (size sample)=	322	RF	1.70E+00	2.37E-01	7.18	45.77	Very strong

Table 6.3 (Continued)								
Model No	**Independent variables**	**Strong Rock Training data**	**Coefficients**		**Standard error**	**Z**	**BIC**	**Grade of significance**
11			Intercept	-7.13708	7.21E-01	-9.90		
	Mix best	2231.91-2277.6=0.90	G	0.09836	0.0765	1.29	-4.12	o
	parameters	940.05-1000.19=0.10	ES	0.0936	1.32E-02	7.08	44.35	Very strong
	n (size sample)=	322	RF	1.70E+00	2.40E-01	7.11	44.75	Very strong
12			Intercept	-7.13708	6.85E-01	-10.43		
	Mix best	2231.91-2277.6=0.90	TDW	1616.6	5354.3	0.30	-5.68	o
	parameters	940.05-1000.19=0.10	ES	0.09724	1.33E-02	7.34	48.08	Very strong
	n (size sample)=	322	RF	-4.11E+02	1.37E+03	-0.30	-5.68	o
13			Intercept	-7.49913	6.85E-01	-10.95		
	Only ES & RF	2231.91-2277.6=0.90	ES	0.09749	0.01327	7.35	48.20	Very strong
		940.05-1000.19=0.10	RF	1.75978	2.38E-01	7.40	49.04	Very strong
	n (size sample)=	322						
14			Intercept	-2.6546	0.45185	-5.87		
	F, G, H	2231.91-2277.6=0.90	F	-73.02527	10.2205	-7.14	45.28	Very strong
	parameters	940.05-1000.19=0.10	G	0.12891	0.07784	1.66	-3.03	o
	n (size sample)=	322	H	-210.1213	26.32945	-7.98	57.91	Very strong
15			Intercept	0.2345	0.1315	1.78		
	ONLY H	2231.91-2277.6=0.90	H	-86.5592	12.0202	-7.20	46.08	Very strong
	parameter	940.05-1000.19=0.10						
	n (size sample)=	322						
16			Intercept	-6.1355	0.8013	-7.66		
	ES, RF, H	2231.91-2277.6=0.90	ES	0.0944	0.0124	7.61	52.18	Very strong
	parameters	940.05-1000.19=0.10	RF	1.3048	0.27983	4.66	15.97	Very strong
	n (size sample)=	322	H	-52.28546	20.48382	-2.55	0.74	Weak

The results of that evaluation are shown in Table 6.4. The Model 3 and the Model 13 have 98% of correct prediction in the training intervals. Model 5 (new parameters) has an overall percentage of correct prediction of 74%, which is low. Model 16 (H, ES and RF) has the highest (99.17%) percentage of correct prediction.

Table 6.4. Evaluation of Logistic Regression Models for Predicting STRONG ROCK in Oklahoma Well													
EVALUATION		**No 3- Normalized Par.**			**No 5-New Parameters**			**No 13- ES and RF**			**No 16- H, ES and RF**		
		PREDICTED		**% CORRECT**	**PREDICTED**		**% CORRECT**	**PREDICTED**		**% CORRECT**	**PREDICTED**		**% CORRECT**
		0	**1**		**0**	**1**		**0**	**1**		**0**	**1**	
OBSERVED	**0**	**209**	**0**	100.00%	**187**	**22**	89.47%	**209**	**0**	100.00%	**209**	**0**	100.00%
	1	**4**	**152**	97.44%	**72**	**84**	53.85%	**5**	**151**	96.79%	**3**	**151**	98.05%
OVERALL		98.90%			74.25%			98.63%			99.17%		

In order to find the best model for strong rock, the four models were applied and evaluated for the entire well (675'-2277'). The response of each model is illustrated in Figure 6.4 with the drilling parameters and the GR plots. As seen in this figure, Model 5 (red curve in top plot) gives an inconclusive diagnosis from 675' to 1,975'. The probability or prediction is about 0.5 during this long interval, which makes it difficult to define if the formation drilled is weak or strong. In addition, this model showed a much lower percentage of correct prediction than models 3, 13 and 16, see Table 6.4.

Model 3 and Model 13, green and blue curves respectively in the top plot, have similar responses. That happens because the only difference between these models is the inclusion of FORS in Model 3, but this parameter has no significance in that model, consequently the models are almost equal. Model 16, the black curve, tends to be equal to Model 3 and 13. That happens because these three models have two common independent variables (Es and RF). However, because of the effect of the new parameter "H", Model 16 shows lower values of probability of strong rock in balling intervals such as 1,500'-1,600' and 1,900'-1,975' making easier to distinguish between strong rock and bit balling situations. In addition, this model responds correctly to known conditions and clearly described the strongest formations drilled during the last bit run.

For all of the above reasons, Model 16 using the normalized parameters ES and RF and the new parameter H as independent variables was selected as the best model for diagnosing strong rock in the Oklahoma Well. It is important to observe that in Chapter 5, for Matagorda Island Well # 1, the best logistic model for strong rock prediction was also obtained when using the normalized parameters as independent variables, where ES and RF had the highest grade of significance. Then, these two models are almost similar because they have common independent variables, ES and RF, with the highest significance in the model.

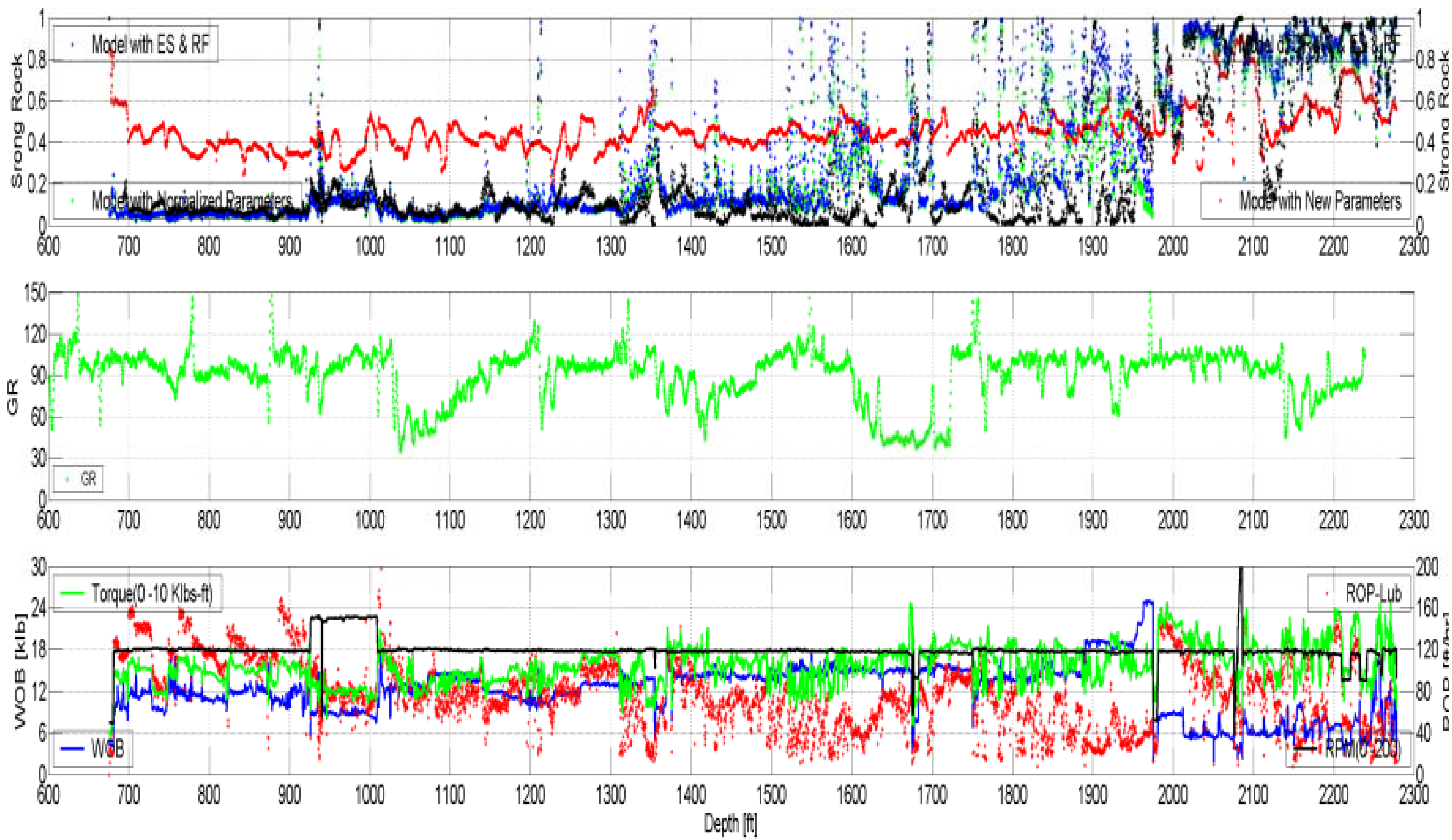

Figure 6.4. Logistic Regression Models for Diagnosing STRONG ROCK in Oklahoma Well.

6.5 Application of Logistic Models in Oklahoma Well.

Once the best logistic models for predicting strong rock and bit balling are selected, they are applied independently. Then, the results are interpreted according to the highest value of probability. In the zones with considerable probability (P>0.5) of both bit balling and strong rock, one of them generally has more marked tendency to the value of probability = 1, so this is considered to be the more likely occurrence. The intervals with low values of probability (P<0.5) for both bit balling and strong rock are considered as "weak rock-clean bit" , specifically weak rocks drilled with a clean bit. Figure 6.5 illustrates the results of applying both models independently to the interval 675'-2,277' of the Oklahoma Well. In this figure, the probability that balling has occurred based on Model 13 from Table 6.1 is plotted at the top in red, and the probability that strong rock is being drilling based on Model 16 from Table 6.3 is shown in the second plot in black. The plot of logs and drilling parameters are also shown. The entire section (675' – 2,277') was divided into the same 23 intervals used when evaluating Aghassi's method. These sections are also shown in this figure.

Table 6.5 describes the results of the logistic diagnosis. For each interval analyzed, the geological interpretation, the probability of bit balling and strong rock, and the logistic diagnosis are shown. Results of probabilities are reported as the average within the interval. As seen in this table, 16 of 23 intervals have correct diagnosis.

Table 6.6 summarizes the evaluation of the results obtained after applying logistic regression models to the Oklahoma Well. The reliability of the method was calculated not only determining the percentage of each of the actual situations that was diagnosed correctly but also determining the percentage of each class of diagnosis that was correct.

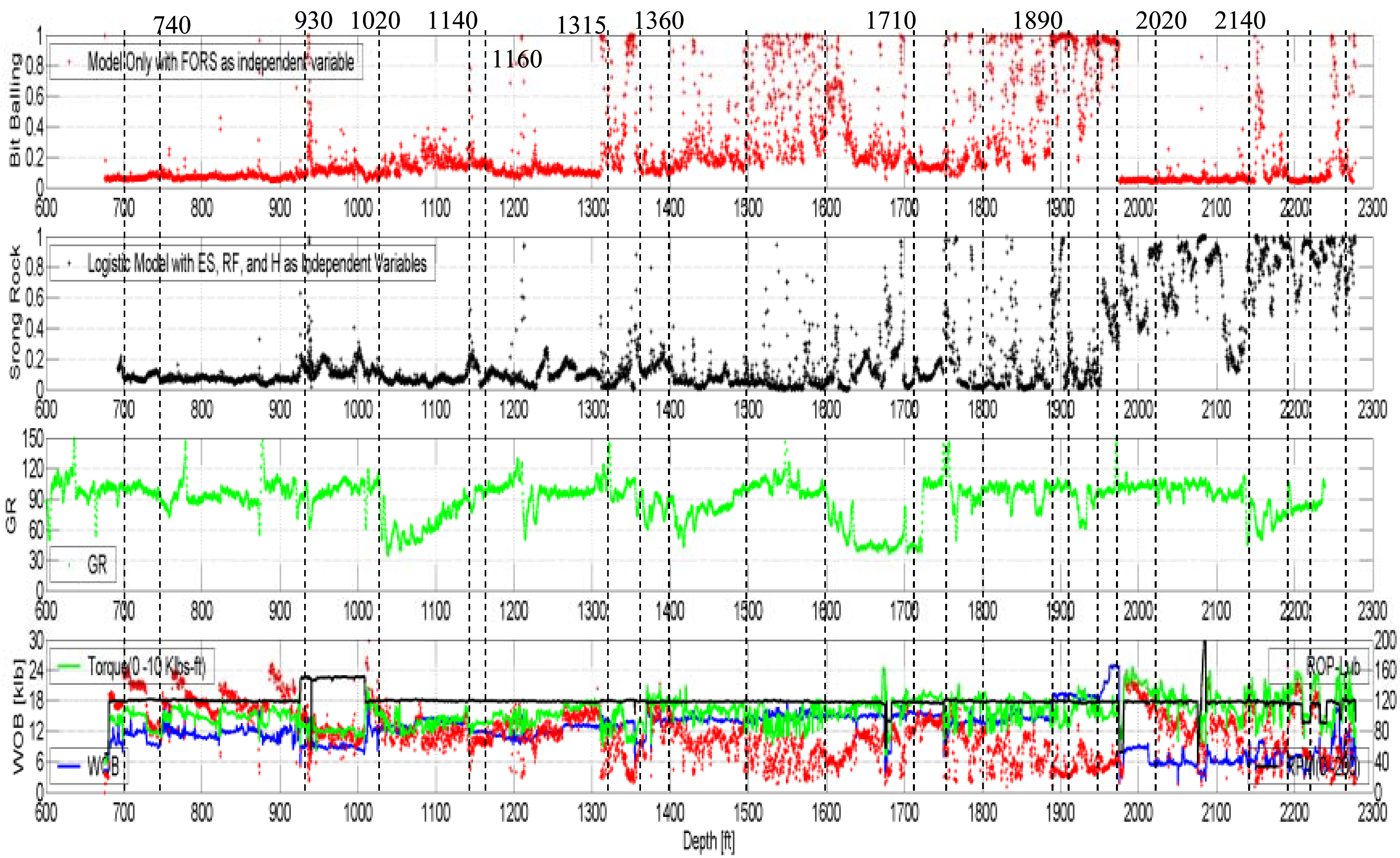

Figure 6.5 Application of Logistic Models in Oklahoma Well.

Table 6.5 Logistic Models Diagnosis - Oklahoma Well					
Interval	**Geologic Interpretation**	**Prob. Balling**	**Prob. Strong Rock**	**Log. Models Diagnosis**	**Correct?**
700'-730'	(Shale-Clean)	0.06	0.07	**Weak Clean Rock**	**YES**
730'-740'	(Shale-Clean)	0.08	0.10	**Weak Clean Rock**	**YES**
740'-940'	(Shale-Clean)	0.11	0.10	**Weak Clean Rock**	**YES**
940'-1,020'	(Shale-Clean)	0.12	0.14	**Weak Clean Rock**	**YES**
1,020'-1,140'	(Sand)	0.16	0.07	**Weak Clean Rock**	**YES**
1,140'-1,160'	(Shale-Balling)	0.20	0.15	**Weak Clean Rock**	**NO**
1,160'-1,315'	(Shale-Clean)	0.15	0.15	**Weak Clean Rock**	**YES**
1,315'-1,360'	(Shale-Clean)	0.48	0.12	**Weak Clean Rock**	**YES**
1,360'-1,400'	(Shaly Sandstone)	0.15	0.17	**Weak Clean Rock**	**YES**
1,400'-1,500'	(Siltstone)	0.28	0.07	**Weak Clean Rock**	**YES**
1,500'-1,600'	(Shale-Balling)	0.55	0.09	**Balling**	**YES**
1,600'-1,710'	(Sand & Siltstone)	0.29	0.19	**Weak Clean Rock**	**YES**
1,710'-1,750'	(Shale-Clean)	0.14	0.12	**Weak Clean Rock**	**YES**
1,750'-1,800'	(Strong Siltstone)	0.35	0.24	**Weak Clean Rock**	**NO**
1,800'-1,890'	(Strong Siltstone)	0.52	0.14	**Balling**	**NO**
1,890'-1,910'	(Shale-Balling)	0.93	0.48	**Balling**	**YES**
1,910'-1,950'	(Shale getting balled)	0.80	0.11	**Balling**	**YES**
1,950'-1,975'	(Severe Balling)	0.92	0.46	**Balling**	**YES**
1,975'-2,020'	(Shale-Clean)	0.05	0.67	**Strong Rock**	**NO**
2,020'-2,140'	(Shale-Clean)	0.07	0.68	**Strong Rock**	**NO**
2,140'-2,195'	(Siltstone)	0.21	0.85	**Strong Rock**	**NO**
2,195'-2,225'	(Shale-Clean)	0.05	0.84	**Strong Rock**	**NO**
2,225'-2,270'	(Strong Sandstone)	0.26	0.84	**Strong Rock**	**YES**

As seen in Table 6.6, the situation that was most likely to be diagnosed correctly by the models was bit balling (80%). Strong rock and weak rock-clean bit were diagnosed correctly 33% and 73% of the time respectively. Weak rock-clean bit were the greater amount (15) of intervals where the models were applied.

In summary, 23 intervals were analyzed with logistic models in the Oklahoma Well, and 16 of these intervals were correctly diagnosed. Consequently, the overall accuracy of logistic models is 16/23 or 70%.

Table 6.6 - Prediction Evaluation - Logistic Models in Oklahoma Well

	DIAGNOSIS					
	Situation	**Bit Balling**	**Strong Rock**	**Weak rock-clean bit**	**TOTAL**	**% of Situations Diagnosed Correctly (Accuracy)**
REAL	**Bit Balling**	**4**	0	1	5	**80%**
	Strong Rock	1	**1**	1	3	**33%**
	Weak-Clean Rock	0	4	**11**	15	**73%**
	TOTAL	5	5	13	23	**70%**
	% Correct of Diagnoses (Reliability)	**80%**	**20%**	**85%**	**70%**	

Considering the reliability of each of the three possible diagnoses, 80% of the bit balling predictions and 85% of the predictions that the bit was clean and drilling weak rock were correct. However, only one strong rock prediction was correct for a 20% reliability. These results also give an overall reliability of 70%, but show that the strong rock diagnoses are very likely to be incorrect and actually to be weak rock with clean bit. However, the wrong diagnoses, 4 of 5, as strong rock when the real situation is weak rock-clean bit would not cause problems during drilling operations because the response of increasing the WOB if more ROP is desired would be acceptable until the onset of balling.

It is important to note that the weak rock-clean bit diagnosis when the real situation is bit balling is very disadvantageous for drilling operations, because if the driller increases WOB in order to improve ROP, the balling situation will be more severe. The preferred response would be opposite to try to minimize the severity of the balling. Although that situation happened only in one of the 13 diagnoses as weak rock-clean bit, this represents a weakness when using logistic models for diagnosing bit performance. It is a significant concern because the negative impact of balling typically continues long after the initial situation.

The interpretation that bit balling is occurring when the real situation is strong rock would likely cause a lower ROP because the reaction of the driller would be to decrease WOB in order to avoid a severe bit-balling situation. The strong rock requires a higher WOB to be drilled faster. However, this problem is not as damaging to bit performance as a severe bit balling because the effect lasts only for the duration of the strong rock interval.

6.6 Comparison of Results from Aghassi's Method and Logistic Regression Models

The results obtained in Oklahoma Well were used to compare the diagnosis obtained from Aghassi's method and the logistic models. Table 6.7 summarizes the results obtained after applying Aghassi's method to the Oklahoma Well. This table has the same information as Table 6.6 for logistic regression models, and has 21 instead of 23 intervals because the baselines are not included in the diagnosis. In addition, in order to make the results from the two procedures equivalent, the situations "Sand" and "Shale Clean Bit" established by Aghassi are considered together as Weak rock-Clean bit.

As seen in Tables 6.6 and 6.7, the situation that was most likely to be diagnosed correctly by Aghassi's method was bit balling (60%). Logistic models diagnosed bit balling with higher accuracy (80%). Strong rock and weak rock-clean bit were diagnosed correctly by Aghassi's method 33% and 23% of the time respectively.

Table 6.7 - Prediction Evaluation - Aghassi Method in Oklahoma Well							
	DIAGNOSIS						
	Situation	**Bit balling**	**Strong Rock**	**Weak rock-Clean bit**	**NO CLEAR DEFINITION**	**TOTAL**	**% of Situations Diagnosed Correctly (Accuracy)**
REAL	**Bit Balling**	**3**	0	0	2	5	**60%**
	Strong Rock	0	**1**	0	2	3	**33%**
	Weak-Clean Rock	0	3	**3**	**7**	13	**23%**
	TOTAL	3	4	3	11	21	**33%**
	% Correct of Diagnoses (Reliability)	**100%**	**25%**	**100%**	**0%**	**33%**	

Considering the reliability of each of the three possible diagnoses, for Aghassi's method 100% of the bit balling diagnoses and 100% of the diagnosis that the bit was clean and drilling weak rock were correct, while these values were 80% and 85% respectively for logistic models. However, both Aghassi's method and logistic models showed a low percentage of correct diagnoses, 25% and 20% respectively, when they diagnosed "strong rock". The overall reliability of Aghassi's method was 33%, which is lower than the 70% obtained by logistic models.

Given these results, neither Aghassi's method nor logistic regression models supplied an acceptable final global diagnosis if used independently for diagnosing bit performance. The reliability of Aghassi's method was drastically affected by the 11 intervals with no clear diagnosis. It is important to note that almost all of these intervals were correctly diagnosed by the logistic model. In addition, the reliability of logistic regression models was affected by the wrong diagnosis of four of the five intervals drilled by bit # 5, which were diagnosed correctly by Aghassi's method. Consequently, it seems that the shortcomings of Aghassi's method may be overcome by logistic models and vice versa. Therefore, in an attempt to improve the final reliability of diagnosing bit performance, I studied the combined application of both procedures by using logistic regression models as a complement to Aghassi's method.

6.7 Logistic Models as a Complement to Aghassi's Method in Oklahoma Well

The technique proposed to improve the final diagnosis is as follows. First, apply Aghassi's method and select the intervals with no clear diagnosis. Then, develop the best logistic models for predicting strong rock and bit balling. After that, these logistic models are applied, only to the intervals with partial or no clear diagnosis from Aghassi's method. Then, a final diagnosis is obtained.

Table 6.8 Logistic Regression as a Complement to Aghassi's Method in OklahomaWell

Interval	Geological and Operational Interpretation	Aghassi's Method Diagnosis			Logistic Regression	
		Conventional Parameters	New Parameters	Is Correct?	Diagnosis	Is Correct?
730'-740'	Shale - Clean Bit	Shale-Clean Bit	Bit Balled	?	Weak rock-Clean bit	YES
1,020-1,140'	Sand	Bit Balled	Sand or Stronger Rock	?	Weak rock-Clean Bit	YES
1,315'-1,360'	Shale - Clean Bit	Bit Balled	Shale - Clean Bit	?	Weak rock-Clean Bit	YES
1,360'-1,400'	Shaly Sandstone	Strong Rock	Bit Balled	?	Weak rock-Clean Bit	YES
1,500'-1,600'	Shale, Bit getting Balled	Bit Balled	Shale - Clean Bit	?	Bit Balled	YES
1,600'-1,710'	Sand & Siltstone	Bit Balled	Shale - Clean Bit	?	Weak rock-Clean Bit	YES
1,800'-1,890'	Strong Siltstone	Bit Balled	Shale - Clean Bit	?	Bit Balled	NO
1,910'-1,950'	Sandstone & Shale Balled	Bit Balled	Shale - Clean Bit	?	Bit Balled	YES
1,975'-2,020'	Shale - Clean Bit	Shale-Clean Bit	Bit Balled	?	Strong Rock	NO
2,195'-2,225'	Shale - Clean Bit	Shale-Clean Bit	Bit Balled	?	Strong Rock	NO
2,225'-2,270'	Strong Sandstone	Strong Rock	Bit Balled	?	Strong Rock	YES

Table 6.8 shows a summary, extracted from Appendix I and Table 6.5 of the results obtained for the 11 intervals with partial or wrong diagnosis from Aghassi's method and the results obtained after applying logistic regression models to these intervals.

After re-evaluating with logistic regression models, the following observations or diagnoses were made for the 11 intervals described in table 6.8.

- 730'-740': In this section, the probability of bit balling and strong rock are 0.08 and 0.10 respectively. Then, because of this low probability of occurrence of these events, this interval has neither bit balling nor strong rock. Then, the final diagnosis is that this is a weak formation and clean bit, which agrees with the lithology interpretation made from logs and core description in this interval.
- 1,020'-1,140: The probability of bit balling and strong rock are 0.16 and 0.07 respectively in this section. Like the previous interval, the low probability of occurrence of these events makes the diagnosis of this interval neither bit balling nor strong rock. Then, the final diagnosis is that this is a weak formation and clean bit, which agrees with the lithology interpretation made from logs and cores.
- 1,315'-1,360'. In this section, there is a relatively high probability of bit balling (0.48) but it is not as high enough to be considered a bit-balling situation. The strong rock prediction is 0.12. Then, the diagnosis is that the formation is weak and the bit may have started to become balled, which agrees with the geological and operational interpretation because at 1,350' the drill string was pulled out to check the bit because the ROP had dropped to 17 ft/hr. However, the bit was found to be in good condition with no obvious explanation for the decrease in performance.

- 1,360'-1,400': This is another section with low probability of strong rock and bit balling, 0.17 and 0.15 respectively. Then, it is diagnosed as a weak formation drilled with clean bit, which is true.
- 1,500'-1,600: As seen in Figure 6.5 and Table 6.5, there is clearly a high probability of bit balling in this long section (0.55 balling versus 0.09 strong rock). Consequently, the final diagnosis is that the bit is becoming balled, which agrees with the operational and geological interpretation, although there is no definite proof that the bit was balled.
- 1,600'-1,710': Although there are reasonable probabilities of bit balling and strong rock in this interval, 0.29 and 019 respectively, it is considered as a weak section drilled with clean bit, which agrees with the geological and operational interpretation.
- 1,800'-1,890': In this section of stronger rock the probability of balling (0.52) is higher than the probability of strong rock (0.14). Consequently, the diagnosis from logistic models is not correct.
- 1,910'-1,950': According to the geological interpretation this is an interval of shale with a sand section interbedded between approximately 1,925' and 1'940'. The logistic diagnosis shows that at the beginning (1,910'-1,925') and at the end (1,940'-1,950') of the interval the bit is balled. The average probabilities of balling and strong rock in this interval are 0.8 and 0.11 respectively, then the final diagnosis is that the bit is balled. Again, the logistic diagnosis is close to what happened in the field, because at 1,940' the bit balling situation began to become severe, which was confirmed when this bit # 4RR, was tripped out and observed to be balled.
- 1,975'-2,020' and 2,195'-2,225': In these two intervals of weak rock with clean bit the logistic diagnosis is "strong rock", which is not correct.

- 2,225’-2,270’: As seen in figure 6.5 and Table 6.5, the diagnosis from logistic models is that a stronger rock was drilled in this interval, which is correct.

6.7.1 Effects and Evaluation of Applying Logistic Models as a Complement to Aghassi’s Method in Oklahoma Well

After applying logistic regression as a complement to Aghassi’s method, 8 of the 11 intervals, which had inconclusive diagnosis, have clear and validated diagnoses. Consequently, the final diagnosis from this complemented technique would be as shown in Table 6.9.

Table 6.9 - Final Diagnosis after Applying Logistic Models as Complement to Aghassi's Method in Oklahoma Well

		DIAGNOSIS				
	Situation	**Bit balling**	**Strong Rock**	**Weak rock-Clean bit**	**TOTAL**	**% of Situations Diagnosed Correctly (Accuracy)**
REAL	**Bit Balling**	**5**	0	0	5	**100%**
	Strong Rock	1	**2**	0	3	**67%**
	Weak rock-Clean bit	0	5	**8**	13	**62%**
	TOTAL	6	7	8	21	**71%**
	% Correct of Diagnoses (Reliability)	**83%**	**29%**	**100%**	**71%**	

As seen in this table, the situation that was most likely to be diagnosed correctly by the complemented technique was bit balling (100%) which is greater than the values obtained by Aghassi’s method (60%) and logistic models (80%). Strong rock was diagnosed correctly in 67% of the cases. This accuracy is greater than that obtained by Aghassi or logistic models alone (33%). Weak rock-clean bit was diagnosed correctly by the complemented technique in 62% of the cases. That value is greater than with Aghassi (23%) but less than with the logistic model (73%). In summary, 21 intervals were analyzed with the complemented technique in the

Oklahoma Well, and 15 of these intervals were correctly diagnosed. Consequently, the overall accuracy of complementary use of the models is 15/21 or 71%.

Considering the reliability of each of the three possible predictions, 83% of the bit balling diagnoses and 100% of the diagnoses that the bit was clean and drilling weak rock were correct using the complemented technique. However, only two strong rock predictions were correct for a 29% reliability. These results also give an overall reliability of 71%, but show that conditions diagnosed as strong rock are very likely to be weak clean rock.

In summary, the global reliability and accuracy of the final diagnosis was improved considerably from 33%, using only Aghassi's method, to 71% using the complemented technique. Although this result is almost the same as obtained when using only logistic models (70%), the complemented technique provides a better accuracy and reliability for each of the three situations predicted. In addition, the shortcoming of predicting weak-rock when the situation is bit balling is eliminated.

6.8 Conclusions

The best logistic model for predicting a balled bit was obtained considering only the normalized parameter FORS as an independent variable. For strong rock prediction, the best logistic model was obtained using the normalized parameters ES and RF together with the new parameter H as independent variables.

The torque parameter, BT, proposed by Aghassi[1] as a means to distinguish if a ROP reduction is due to a balled bit or strong rock, was tested in logistic models as an independent variable for both strong rock and bit balling predictions. The results showed that this parameter has no statistical significance in the models.

Logistic models supplied a wrong weak rock-clean bit rock prediction when the real situation was bit balling which is highly undesirable, because it can cause a severe balling situation. This represents a disadvantage when using logistic models for predicting bit performance.

According to the results obtained from the example illustrated in this chapter, the final diagnosis of bit performance can be improved from 33 %, when applying only Aghassi's method, to 71%, when applying logistic models as a complement to Aghassi's method. Although this result is almost the same as obtained when using only logistic models (70%), I consider the complemented technique to be the best way to diagnose bit performance because it provides a better accuracy and reliability for each of the three situations predicted. In addition, the shortcoming of predicting weak-rock when the situation is bit balling is eliminated.

7 SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

7.1 Summary

Aghassi and Smith's method[1,37] was evaluated comparing MWD and surface data for a given interval of a well, and the accuracy and utility of using surface data instead of data measured at the bit were assessed.

The method was also applied using surface data in four additional wells with known lithology and changes in bit performance. The reliability of the method was evaluated.

Statistical logistic models were studied and proposed as a means to diagnose bit performance, especially for distinguishing between strong rock and a balled bit. The models were evaluated according to the percentage of correct prediction. These logistic models were also applied as a complement to Aghassi and Smith's method, and the impact on the final diagnosis was evaluated.

7.2 Conclusions

1) The results obtained when applying Aghassi and Smith's method for diagnosing bit performance using data measured at the surface are almost identical to results obtained using down-hole MWD data. Therefore surface data is concluded to be a practical basis for diagnosing bit performance.
2) The interpretation of Aghassi and Smith's diagnostic parameters when drilling strong rock is often incorrect because the diagnosis is that the bit is balled. One of the apparent situations that causes this error is a lack of variation in the torque measured at surface when changing from a weak formation to strong rock. Apparently, that happens in the field when the driller concluded that a strong rock was encountered and decreased RPM and increased WOB as the ROP decreased. If torque does not increase as expected in strong rock, the interpretation is

that bit balling is occurring. Therefore, the response of the two parameters, “Torque/WOB” and “F” which distinguish between strong rock and bit balling, is different than expected.

3) The accuracy of the diagnosis when applying Aghassi and Smith’s method is sometimes dependent on the selection of the baseline zone. The value of the WOB is apparently important in the selection of the baseline, because it affects the diagnosis of the parameters “Torque/WOB” and “F”. Specifically if the baseline is set in an interval with relatively low WOB, the Torque/WOB in zones of stronger rock will be smaller or lower than its baseline value because the increase in WOB will be proportionately higher than the increase in Torque when drilling the stronger rock. The resulting diagnosis is a balled bit instead of strong rock. On the other hand, if the baseline is in a section drilled with relatively high WOB, the Torque/WOB is likely to be slightly larger or close to its value in the baseline, as required for a correct diagnosis of strong rock. Therefore in addition to locating the baseline in a shale interval with relatively high ROP, the baseline should preferably be set in an interval of relatively high WOB.
4) The overall accuracy of Aghassi and Smith’s method, calculated after applying the method in five wells, was 51 %. The analysis also shows that the highest accuracy for a well was 67% and the lowest was 33%. The accuracy of the method was calculated not only for each well, but also for each of the situations or lithology diagnosed by the method: sand, shale with clean bit, shale with bit balling, and strong rock. As seen in Table 3.12, the situation that was most likely to be diagnosed correctly was bit balling (82%). Strong rock, shale-clean bit and sand were diagnosed correctly 51%, 40% and 50% of the time respectively.
5) Logistic regression models were developed and tested for diagnosing bit balling and strong rock. Values greater than 90% were obtained when the models were evaluated against training intervals using the criterion of percent of correct prediction. Applying logistic models for

diagnosing bit balling and strong rock individually was tested as a means to determine which of these two situations is more likely to be occurring. Based on the analysis of two examples of applying the combination of these regression models, they are potentially useful for distinguishing between strong rock and a balled bit. In addition, no baseline is necessary for these models. However, the models do require that known occurrences of balling and strong rock be used as training data for the algorithm. Moreover, from the result obtained in one well, logistic models supplied a wrong weak rock-clean bit rock prediction when the real situation was bit balling which is highly undesirable, because it can result in actions leading to a more severe balling situation. This represents a disadvantage when using logistic models for predicting bit performance

6) From the results of applying Aghassi and Smith's method and logistic regression models independently, it was concluded that neither Aghassi's method nor logistic regression models supplied a fully acceptable final global diagnosis if used independently. However, they provided a better final diagnosis when logistic regression models were applied as complement to Aghassi and Smith's method.
7) Although the best diagnostic parameters (from BIC evaluation) were almost always the normalized parameters (FORS, RF, and ES), the final models for bit balling and strong rock in each well do not have the same independent variables. In other words, there are not conclusive results about which parameters are the best for diagnosing strong rock and bit balling with logistic regression models.
8) The logistic models can be used as a complement to Aghassi and Smith's method in order to improve the reliability and accuracy of the diagnosis. Based on the results obtained from one well, the final diagnosis of bit performance can be improved from 33 %, when applying only Aghassi's method, to 71%, when applying both Aghassi's method and complementary logistic

models. Although this result is almost the same as obtained when using only logistic models (70%), I consider the complemented technique to be the best way to diagnose bit performance because it provides a better accuracy and reliability for each of the three situations predicted. In addition, the shortcoming of predicting weak-rock when the situation is bit balling is eliminated.

7.3 Recommendations

1) Aghassi and Smith's method should be applied to several additional drilling data sets with high resolution versus time, at least one record every one second. The results and final reliability and accuracy of the model should then be re-evaluated. Also, logistic models should be developed for and applied to these data sets, and their impact on final diagnosis evaluated.

2) Another statistical tool that can be used to distinguish between strong rocks and bit balling is the "multinomial logistic regression model" explained in Appendix II. These models extend the analysis of binary variables to the analysis of dependent variables with more than two categories. An example of a dependent variable with more than two categories can be "bit performance" because there are several situations when drilling a formation in a well such as drilling a strong rock, drilling a weak rock with a balled bit, and drilling a weak rock with a clean bit. Then, with the respective training data, a multinomial logistic regression model can be developed for predicting which of these three situations is more likely to occurs. Consequently, only a model would be used to distinguish directly between strong rock and bit balling avoiding the process of developing two independent models and comparing probabilities.

3) If the results promise practical application, Aghassi and Smith's method and logistic models should be programmed in computers to be used in real time in the field. Using MatLab is recommended because it is easy to program the simple calculations of Aghassi's method and logistic regression and it provides a good graphical description of the results, which should be useful for evaluating the methods in the field.

4) The computer-based method should then be evaluated in the field for response to formation change and to the intentional change of drilling parameters. The final reliability of the models and method should be evaluated and the potential practical commercialization assessed.

REFERENCES

1. Arash Aghassi, Investigation of Qualitative Methods for Diagnosis of Poor Bit Performance Using Surface Drilling Parameters, Thesis, Louisiana State University, May 2003

2. John Rogers Smith, "Addressing the Problem of PDC Bit Performance in Deep Shales," IADC/SPE 47814, IADC/SPE Asia Pacific Drilling Conference, Jakarta, Indonesia, September 7-9, 1998

3. C.A. Cheatham and J.J. Nahm, "Bit Balling in Water-Reactive Shale During Full-Scale Drilling Rate Test," IADC/SPE 19926, IADC/SPE Conference, Houston, TX, February 1990

4. I.G. Falconer, T.M. Burgess, and M.C. Sheppard, "Separating Bit and Lithology Effects from Drilling Mechanics Data," IADC/SPE 17191, Drilling Conference, Dallas, TX, February 1988

5. Matthew Bible, Marc Lesage, and Ian Falconer, Method for Detecting Drilling Events from Measurement While Drilling Sensors, USA Patent 4876886, Anadrill Inc., 1989

6. John Rogers Smith, Drilling Over-Pressured Shales with PDC Bits: A Study of Rock Characteristics and Field Experience Offshore Texas, Thesis, Louisiana State University, December 1995

7. John Rogers Smith and Jeffrey Bruce Lund, "Single Cutter Tests Demonstrate Cause of Poor PDC Bit Performance in Deep Shales," ETCE 2000, ASME - Drilling Technology Symposium, Houston, TX, February 14-17, 2000

8. R.C. Pessier, and M.J. Fear, "Quantify Common Drilling Problems with Mechanical Specific Energy and a Bit-Specific Coefficient of Sliding Friction," SPE 24584, 67th Annual Technical Conference and Exhibition of Society of Petroleum Engineers, Washington, DC, October 4-7, 1992

9. John Rogers Smith, Diagnosis of Poor PDC Bit Performance in Deep Shales, Dissertation, Louisiana State University, August 1998

10. John Rogers Smith, "Performance Analysis of Deep PDC Bits Runs in Water-Base Muds," ETCE 2000, ASME - Drilling Technology Symposium, Houston, TX, February 14-17, 2000

11. Arthur Lubinski, Instantaneous Bit Rate of Drilling Meters, USA Patent 2688871, 1949

12. M.J. Fear, "How to Improve Rate of Penetration in Field Operations," IADC/SPE 35107, IADC/SPE Drilling Conference, New Orleans, LA, March 12-15, 1996

13. Charles H. King, Mitchell D. Pinckard, and Jimmy L. Puckett, Method of and System for Increasing Drilling Efficiency, USA Patent 6233498, Noble Drilling Services Inc., 2001

14. Ergun Kuru, Effects of Rock/Cutter Friction on PDC Bit Drilling Performance, An Experimental and Theoretical Study, Dissertation, Louisiana State University, May 1990

15. Adam T. Bourgoyne Jr., Martin E. Chenevert, Keith K. Milheim, and F.S. Young Jr., Applied Drilling Engineering, SPE Text Book Series, Richardson, TX, 1991

16. Mitchell D. Pinckard, Method of and System for Optimizing Rate of Penetration in Drilling Operations, USA Patent 6192998, Noble Drilling Services Inc., 2001

17. B.G. Chesser and A. C. Perricone, "A Physicochemical Approach to the Prevention of Balling of Gumbo Shale," SPE 4515, 48th Annual Fall Meeting of the Society of Petroleum Engineers of AIME, Las Vegas, NV, September 1973

18. G.A. Cooper and Sanjit Roy, "Prevention of Bit balling by Electro-Osmosis," SPE 27882, SPE Western Regional Meeting, Long Beach, CA, March 23-25, 1994

19. Lee Smith, F.K. Mody, Arthur Hale, and Nils Romslo, "Successful Field Application of An Electro-Negative 'Coating' to Reduce Bit Balling Tendencies in Water Base Mud," IADC/SPE 35110, IADC/SPE Drilling Conference, New Orleans, LA, March 12-15, 1996

20. P.R. Hariharan, G.A. Cooper, and A.H. Hale, "Bit Balling Reduction by Electro-Osmosis While Drilling Shale Using A Model BHA," IADC/SPE 39311, IADC/SPE Drilling Conference, Dallas, TX, March 3-6, 1998

21. E. van Oort, J.E. Friedheim, and B. Toups, "Drilling Faster with Water-Base Muds," AADE, AADE Annual Technical Forum – Improvements in Drilling Fluids Technology, Houston, Texas, March 30-31, 1999

22. Ron Bland, Bill Haliday, Roland Illerhaus, Mat Isbell, Scott McDonald, and Rolf Pessier, "Drilling Fluid and Bit Enhancement for Drilling Shales," AADE, AADE Annual Technical Forum – Improvements in Drilling Fluids Technology, Houston, Texas, March 30-31, 1999

23. Dennis E. O'Brien and Martin E. Chenevert, "Stabilizing Sensitive Shale with Inhibited, Potassium-Base Drilling Fluids," SPE 4232, Journal of Petroleum Technology, Austin, TX, September 1973, p. 1089-1100

24. C.A. Cheatham, J.J. Nahm, and N.D. Heikamp, "Effect of Selected Mud Properties on Rate of Penetration – Full-Scale Shale Drilling Simulations," SPE/IADC 13465, SPE/IADC Drilling Conference, New Orleans, LA, March 6-8, 1985

25. T.M. Warren, "Penetration Rate Performance of Roller-Cone Bits," SPE 13259, SPE Drilling Engineering, Houston, TX, March 1987, p. 9-18

26. T.M. Burgess and W.G. Lesso Jr., "Measuring the Wear of Milled Tooth Bits Using MWD Torque and Weight-On-Bit," SPE/IADC 13475, SPE/IADC Drilling Conference, New Orleans, LA, March 6-8, 1985

27. L.W. Ledgerwood III and D.P. Salisbury, "Bit Balling and Well-Bore Instability of Down-Hole Shale," SPE 22578, 66th Annual Technical Conference and Exhibition of Society of Petroleum Engineers, Dallas, TX, October 6-9, 1991

28. E. van Oort, R Bland, R. Pessier, "Drilling More Stable Wells Faster and Cheaper with PDC Bits and Water Based Muds," IADC/ SPE paper 59192, presented at the 2000 IADC/SPE Drilling Conference, New Orleans, Feb. 23-25

29. M. Rujhan, M.Z. Bin Zakaria, S. Radford, and D. Eckstrom, "Innovative Low-Friction Coating Reduces PDC Balling and Doubles ROP Drilling Shales with WBM," IADC/ SPE paper 74514, presented at the 2002 IADC/SPE Drilling Conference, Dallas, Feb. 26-28

30. T.M. Warren and W.K. Armagost, "Laboratory Drilling Performance of PDC Bits," SPE 15617, 61st Annual Technical Conference and Exhibition of Society of Petroleum Engineers, New Orleans, LA, October 5-8, 1986

31. Steven Talor, Alain Besson, Djoko Minto, I.S. Mampuk, "Unique PDC Bit Technologies Combine to Reduce Drilling Time in Interbedded Formations," SPE 64005, SPE Drill. & Completion Vol. 15, No. 2, Jakarta, Indonesia, June 2000, p. 112-117

32. B.H Walker, A.D. Black, W.P. Klauber, T. Little, and M. Khodaverdian, "Roller-Bit Penetration Rate Response as a Function of Rock Properties and Well Depth," SPE 15620, 61st Annual Technical Conference and Exhibition Of Society of Petroleum Engineers, New Orleans, LA, October 5-8, 1986

33. Pampel F.C, Logistic Regression, a Primer, Sage University Paper Series on Quantitative Applications in the Social Sciences, series No 07-132. Thousand Oaks, CA: Sage.

34. Menard, Scott, Applied Logistic Regression Analysis, Sage University Paper Series on Quantitative Applications in the Social Sciences, series No 07-106. Thousand Oaks, CA: Sage.

35. J.L Jensen, L.W. Lake, P.W.M. Corbett and D.J. Goggin, Statistic for Petroleum Engineers and Geoscientists, Handbook of Petroleum Exploration and Production 2. Second Edition. 2002

36. A.J. Dobson, An Introduction to Generalized Linear Models; Second Edition. 2002

37. A. Aghassi, J.R Smith, Confidential Disclosure Document, Louisiana State University, October 2002.

38. S. C. Chapra, R.P.Canale, Numerical Methods for Engineers Mc Graw-Hill International Editions, Second Edition, 1990.

39. J.R Smith Confidential Disclosure Document, Louisiana State University, October 2000.

APPENDIX I

DIAGNOSIS FROM AGHASSI'S METHOD – OKLAHOMA WELL

Diagnosis from Aghassi's Method - Oklahoma Well.							
Interval (Logging interpretation)	Conventional Diagnostic Parameters			New Diagnostic Parameters		Interpretation	Is it correct?
	Torque/WOB	Torque/ROP	ROP/WOB	F (Torque, WOB)	G (ROP, WOB)		
700'-730' (Shale)	Baseline-1	Baseline-1	Baseline-1	Baseline-1	Baseline-1	Shale Base-line-1	
730'-740' (Shale-Clean)	Smaller	Same	Smaller	Smaller negative	Noisy	No Conclusive	?
740'-940' (Shale-Clean)	Same	Same	Same	Larger negative	Larger negative	Shale Clean Bit	YES
940'-1'020' (Shale-Clean)	Slightly larger	Slightly larger	Smaller	Larger negative	Larger negative	Stronger Rock	NO
1,020'-1,140' (Sand)	Smaller	Slightly larger	Smaller	Larger negative	Larger negative	Conv: Balling New: Clean	?
1,140'-1,160' (Shale-Balling)	Smaller	Slightly larger	Smaller	Smaller negative	Noisy	BALLING	YES
1,160'-1,315' (Shale-Clean)	Slightly larger	Slightly larger	Smaller	Larger negative	Larger negative	Stronger Rock	NO
1,315'-1,360' (Shale-Clean)	Smaller	Slightly larger	Smaller	Larger negative	Larger negative	Conv: Balling New: Clean	?
1,360'-1,400' (Shaly Sandstone)	Slightly larger	Larger	Smaller	Slightly Smaller negative	Noisy	Conv: Stronger New: Balled	?
1,400'-1,500' (Siltstone)	Close	Larger	Smaller	Larger negative	Larger negative	Stronger Rock	NO
1,500'-1,600' (Shale-Balling)	Smaller	Larger	Smaller	Larger negative	Slightly larger Neg.	Conv: Balling New: Clean	?
1,600'-1,710' (Sand & Siltstone)	Smaller	Larger	Smaller	Same	Noisy	Conv: Balling New: Clean	?
1,710'-1,750' (Shale)	Baseline-2	Baseline-2	Baseline-2	Baseline-2	Baseline-2		
1,750'-1,800' (Strong Siltstone)	Close	Slightly Larger	Slightly Smaller	Same	Same	Stronger Rock	YES
1,800'-1,890' (Strong Siltstone)	Slightly Smaller	Larger	Slightly Smaller	Same	Same	Conv: Balling New: Clean	?
1,890'-1,910' (Shale-Balling)	Smaller	Larger	Smaller	Smaller negative	Smaller Neg.	BALLING	YES
1,910'-1,950' (Shale getting balled)	Smaller	Larger	Smaller	Same	Same	Conv: Balling New: Clean	?
1,950'-1,975' (Severe Balling)	Smaller	Larger	Smaller	Smaller negative	Smaller Neg.	BALLING	YES
1,975'-2,020' (Shale-Clean)	Larger	Sligthly Smaller	Larger	Smaller negative	Smaller Neg.	Conv: Clean New: Balling	?
2,020'-2,140' (Shale-Clean)	Larger	Close	Larger	Larger negative	Larger negative	Shale Clean Bit	YES
2,140'-2,195' (Siltstone)	Larger	Larger	Larger	Slightly larger Neg.	Slightly larger Neg	Sand	YES
2,195'-2,225' (Shale-Clean)	Larger	Sligthly Smaller	Larger	Smaller negative	Noisy	Conv: Clean New: Balling	?
2,225'-2,270' (Strong Sandstone)	Slightly larger	Larger	Varies	Smaller negative	Noisy	Conv: Strong New: Balling	?

APPENDIX II

MULTINOMIAL LOGISTIC REGRESSION MODELS[34]

Logistic regression analysis may be extended beyond the analysis of dichotomous variables to the analysis of dependent variables with more than two categories. In the literature on logistic regression, the resulting models have been called polytomous or multinomial logistic regression models.

Mathematically, the extension of the dichotomous logistic regression model to polytomous dependent variables is straightforward. One value, typically the first or last, of the dependent variable is designated as the reference category, Y = h_0, and the probability of membership in other categories is compared to the probability of membership in the reference category.

For dependent variables with some number of categories *M*, this requires the calculation of *M*-1 equations, one for each category relative to the reference category, to describe the relationship between the dependent variable and the independent variables. For each category of the dependent variable except the reference category, the following equation can be written

$$j_h(X_1, X_2, \ldots\ldots X_k) = e^{(\alpha_h + \beta_{h1} X_1 + \beta_{h2} X_2 + \ldots.. \beta_{hk} X_k)}$$

$h = 1, 2, \ldots\ldots.., M\text{-}1,$

Where the subscript *k* refers to specific independent variables *X* and the subscript *h* refers to specific values of the dependent variable Y. The probability that Y is equal to any value *h* other than the excluded value h_o is

$$P(Y = h / X_1, X_2, \ldots\ldots X_k) = \frac{e^{(\alpha_h + \beta_{h1} X_1 + \beta_{h2} X_2 + \ldots\ldots \beta_{hk} X_k)}}{1 + \sum_{h=1}^{M-1} e^{(\alpha_h + \beta_{h1} X_1 + \beta_{h2} X_2 + \ldots\ldots \beta_{hk} X_k)}}$$

And for the excluded category $h_0 = M$ or 0,

$$P(Y = h_o / X_1, X_2, \ldots\ldots X_k) = \frac{1}{1 + \sum_{h=1}^{M-1} e^{(\alpha_h + \beta_{h1} X_1 + \beta_{h2} X_2 + \ldots\ldots \beta_{hk} X_k)}}$$

The sum of the probabilities must be one,

$$\sum_{h=1}^{M} P(Y = h) = 1$$

Note that when $M = 2$, we have the logistic regression model for the dichotomous dependent variable, the reference category is the first category, $h_0 = 0$, and we have a total of M - 1=1 equations to describe the relationship.

Once P(Y=1) to P(Y=*M*) have been calculated, the highest value of probability defines which of the categories M is more likely to occur.

VITA

Jaime Solano was born to Jaime Solano and Alejandrina Castro in Bucaramanga, Colombia, South America, on February 3, 1966. Jaime received his primary and secondary education in various private and public schools in several cities in Colombia. He subsequently graduated from the Universidad Industrial de Santander, in Bucaramanga, with two Bachelor of Science degrees in civil engineering (June, 1989) and petroleum engineering (September, 1992). Jaime worked for one year at Beltran Pinzon Co., in Bucaramanga, Colombia, as civil engineer, supervising construction of buildings. He later became a drilling engineer and a drilling supervisor working for Ecopetrol, the Colombian national oil company, since 1993